MASTER THESIS

Influence of interaction markers design on player's immersion into a Virtual Reality game narrative.

Author:

Paulina Cywoniuk

Supervisor:

MSc Linda Effinger, MA Silke Hassreiter

This thesis is submitted in fulfilment of the requirements

for the degree of Master of Science in Game Technology

In the programme:

Professional Master Game Technology

Academy for AI, Games and Media

Breda University of Applied Sciences

12/06/2026

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

DECLARATION OF AUTHORSHIP

I, Paulina Cywoniuk, declare that this thesis titled, "Influence of interaction markers design on player's immersion into a Virtual Reality game narration." and the work presented are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Paulina Cywoniuk

Date:

24/09/2025

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

"Short personal statement (typically 2-3 lines, optional) about what you feel you have gained during this study"

Paulina Cywoniuk

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

ABSTRACT

With the increasing availability and affordability of immersive virtual reality hardware, designers face a distinct interface challenge: unlike flat-screen games that rely on 2D overlays, VR demands interfaces that operate within three-dimensional space and ideally support, rather than disrupt, the player's narrative engagement. While prior research has examined diegetic versus non-diegetic UI in VR shooters and spatial games, the influence of UI marker design on narrative immersion specifically — as opposed to general immersion or sense of presence — remains underexplored. This paper addresses that gap by investigating the research question: what is the influence of interaction marker design on players' ability to become immersed in the narration of VR games? Following the diegetics, semiotics, and affordance theories, the researcher developed a VR experience titled "Echoes of Abandonment", in which interactable objects in a digital room were marked according to three experimental conditions — diegetic markers (a brightness shader), non-diegetic markers (an exclamation mark icon), and a no-marker control. The methodology was repeated among 35 participants randomly assigned across the three conditions, with narrative immersion measured through two validated questionnaires — the IEQ-SF and NES — supplemented by custom Likert-scale items designed for the purpose of this study. Quantitative and qualitative analysis revealed no statistically significant differences across conditions, supporting the null hypothesis. These findings contribute to VR interface research by suggesting that marker design may matter less when narrative and UI operate on separate sensory channels, opening directions for further investigation.

Structure of abstract (250 words, one paragraph, no citations)

1. 1 sentece per eah chapter that I've written
2. What is the specific gap that this thsis adresses
3. Paraphrase the research question
4. Methodology (sample size and strudy design)
5. Key findings
6. Conclusion

Abstract draft:

This paper investigates the influence of the design of the interaction marker in the VR experiences on players' ability to become immersed in the narration of the game. Following the diegetics, semiotics and affordance theories, the researcher created a VR game "Echoes of Abandonment" with markers of the interactable objects present in the digital room, that are

used in the experiment. Methodology was repeated among 35 participants, divided into three conditions – diegetic markers, non-diegetic, and the control group. Narration immersion was measured with two questionnaires – IEQ-SF and NES, and custom questions designed by the researcher for the purpose of this study. Analysis included qualitative and quantitative data, that pointed towards the null hypothesis, meaning there is no statistical influence across conditions on players' narration immersion.

Academy for AI, Games and Media

Master of Game Technology

Influence of User Interface markers design on player's immersion into a Virtual Reality game narration

By Paulina Cywoniuk

The purpose of this paper is to explore the influence of the markers design on player's immersion into a story told in a VR game.

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

ACKNOWLEDGEMENTS

The acknowledgments and the people to thank go here, don't forget to include your project advisors

CONTENTS

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

LIST OF FIGURES

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

LIST OF TABLES

1 LIST OF EQUATIONS

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

LIST OF ABBREVIATIONS

UI	User Interface
VR	Virtual Reality
UX	User Experience
6DoF	Six degrees of freedom
HCI	Human-Computer Interaction
IEQ	Immersion Experience Questionnaire (Jannet et al. 2028)
IEQ-SF	Immersion Experience Questionnaire (Cutting et al. 2025)
NES	Narration Engagement Scale ()

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

For/Dedicated to/To my...

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

1 INTRODUCTION

This thesis aims to research the influence of the design of markers on tangible user interface on player's immersion into the narrative of a Virtual Reality (VR) game. The researcher will do so, by testing in a controlled environment of a self-made VR game and with separate test groups different markers design on player's immersion following a pre-made, validated questionary.

The focus of this research is concerned with the VR games due to their increasing relevance in both private and commercial usage. In the recent years market researchers observed a growing interest in the VR technology[1]. Its usage can be spotted in both entertainment games and serious games that we can experience in museums, as training at work, or study. This is in part thanks to the technological advancement in the recent years that lowered the price of an average VR headset, making it more affordable for companies and everyday users (from $55 000[2] back in 2007 to $300[3] in 2025), while simultaneously improving display quality. As a result, the number of active VR users worldwide is expected to reach 216 million by the end of 2025[4].

This growing accessibility makes VR an increasingly important platform for interactive storytelling - through spatial audio, embodied interaction, and visual immersion, VR offers unique possibilities for narrative experiences that influence how people communicate, learn, and think (Irshad & Perkis, 2020). Player's engagement with the story told in a VR game will be one of the focus points of this thesis. This comes from a personal motivation and interest in storytelling in the interactive media, as can be proven by researcher's bachelor's thesis "Storytelling in interactive environments"[5].

---

[1] Vardomatski, S. (2025, September 16). Top 10 VR Trends of 2025: Future of Virtual Reality. *HQSoftware*. https://hqsoftwarelab.com/blog/virtual-reality-trends/

[2] Peter C. Ashline & Vincent S. Lai (1995) VIRTUAL REALITY An Emerging User-Interface Technology, Information Systems Management, 12:1, 82-85, DOI: 10.1080/07399019508962962

[3] *Meta Quest 3*(n.d.). https://www.walmart.com/browse/video-games/meta-quest-3/2636_5454301_3444947_9326049_5989670

[4] Markets, R. A. (2022, April 21). By 2025, the Worldwide User Base of AR and VR Games is Expected to Increase to 216 Million Users. *GlobeNewswire News Room*. https://www.globenewswire.com/news-release/2022/04/21/2426130/28124/en/By-2025-the-Worldwide-User-Base-of-AR-and-VR-Games-is-Expected-to-Increase-to-216-Million

[5] Cywoniuk, P. (n.d.). *Paulina thesis*. https://kabk.github.io/go-theses-24-paulina-cywoniuk/

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

However, these same qualities that make VR compelling for storytelling also create new design challenges, particularly for user interfaces that must function within immersive 3D environments rather than on flat screens.

With the increasing availability, quality, and affordability of immersive gaming and virtual/augmented/mixed reality hard- and software, there is an increasing need for improved interfaces that combine high usability and learnability with a more embodied interaction[6]. Designing user interfaces for VR presents distinct challenges. Unlike flat-screen games that rely on 2D overlays such as menus and button prompts, VR environments demand interfaces that function within three-dimensional space.
L. Bouwels (personal communication, October 1, 2025) argued that UI in VR can be textual, or symbolic. Examples of textual UI in VR could be subtitles, and symbolic UI elements could be any object in the 3D world with which plyers can interact and as a result receive output. This paper addresses the symbolic interface. From a Human-Computer Interaction (HCI) perspective, 3D objects that are manipulated directly by the user mediate the communication between user and the system, hence can be considered UI element[7] (Hepperle et al., 2019). In VR, diegetic UI merges player, world, and interface, where the world becomes the interface and the player becomes the controller[8] (see if this is not a quote + add source). This creates an opportunity unique to VR - rather than separating gameplay from interface, designers can create UI where the world becomes the interface and interaction with objects supports both functionality and narrative coherence.

One of the major components of UI are "markers" that signalize with what object in the 3D digital space the player can interact. In traditional screen-based games, players can discover interactable objects through fast experimentation, such as hovering a cursor, pressing buttons, or looking for UI prompts. In VR, however, such experimentation might be difficult due to the spatial design of the virtual space. The absence of clear interaction cues forces players to break from the narrative experience to search, grab, and test objects through trial and error, potentially disrupting the very immersion that makes VR compelling as a storytelling medium. Markers are the visual indicators that communicate which objects can be manipulated - therefore become essential not merely for usability, but for sustaining narrative engagement. There seems to be a research gap of how to the design markers for the objects in VR experiences that don't break the immersion of the story that the player is experiencing. This thesis will

---

[6] Riecke, B. E., LaViola, J. J., & Kruijff, E. (2018, August 1). *3D user interfaces for virtual reality and games*. https://doi.org/10.1145/3214834.3214869

[7] Wikipedia contributors. (2025a, September 1). *3D human–computer interaction*. Wikipedia. https://en.wikipedia.org/wiki/3D_human%E2%80%93computer_interaction

[8] Archiact. (2021, June 29). Using Diegetic UI in VR games? *Archiact*. https://www.archiact.com/post/uncharted-territory-3-diegetic-ui-in-vr-games

attempt to research this relationship and provide advice for the future researchers and designers.

The research of this thesis will be conducted on a case study game made by the author in her Bachelor's. The VR experience titled "Echoes of Abandonment" tells a story of a child who is dealing with their parents' divorce. By putting on VR goggles players are put in a world sculpted by hand in clay to resemble a kid's bedroom. Behind the door they hear parents arguing. The players goal, as a child, is to mute this argument by playing with toys present in the room. Each interactable toy has a sound attached to it, so that when they keep on interacting with the toys, they create a musical composition that allows them to distance themselves from the parents' fight.

By marking which toys players can interact with, this research tests how specific UI design decisions influence narrative immersion in VR games. Toys, in this case, are 3D objects, and UI elements of the VR experience. Some of the toys in this virtual environment are interactable, and some are not – players can play with toys but not move furniture. By marking the interactable toys, I wish to test how specific UI design decisions influence players' narrative immersion in VR games. Through [my methodology - experiment], this research will provide evidence-based guidelines for designing UI markers of the 3D objects that enhance rather than disrupt story immersion in VR environments. The central question this thesis addresses is: **What is the influence of different markers design in VR games on the player's ability to become immersed in narrative content among young adults living in the Netherlands?**

To investigate this relationship between marker design and narrative immersion, this thesis draws on three theoretical frameworks: diegetic theory, which categorises UI elements by their integration into the game world; affordance theory and semiotics, which explains how objects communicate their interactive possibilities; and narration engagement theory, which describes how narrative absorption can be maintained or disrupted. Together, these frameworks provide the analytical tools to examine how UI design choices affect not just usability, but sustained engagement with story without causing cognitive overload.

2 LITERATURE REVIEW

This section of the thesis will dive into the existing research on the influence of markers of interactive objects (pick-up interaction) in VR games and how that effects player's sense of immersion in the game story. Important concepts will be explained to understand the theoretical base of this research, such as the Diegetic Theory (Fagerholt and Lorentzon, 2009), Affordance (Gibson, 1979), and Narration Immersion (Haggis-Burridge, 2020; Green and Brock's, 2000). Visual feedback on hover

2.1 DIEGETIC THEORY

User interface (UI) is the medium through which the communication between users and computers takes place (Hix and Hartson, 1993). In the game's design industry, UI has been implemented for the player to either interact with the game world or to present information relevant for the game such as dialogue boxes, markers, or menus. There are four types of UI

defined in the Diegetics theory by Fagerholt and Lorentzon (2009) – diegetic, non-diegetic, spatial and meta. The figure below presents those types of UI on the real-world examples from video games.

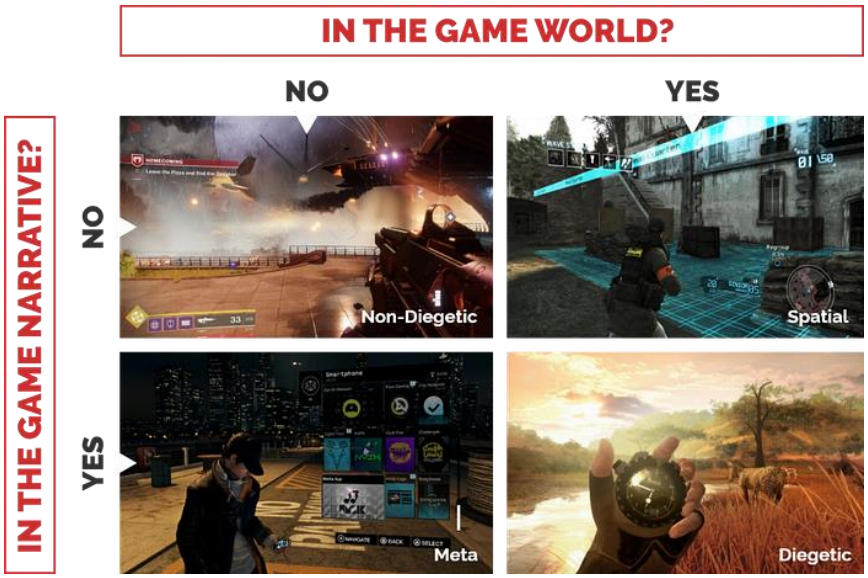

Figure 2.1 – Games (Left to right, top to bottom): Destiny 2, Ghost Recon Future Soldier, Watch Dogs 2, Far Cry 2. source: Kraj, N. (2020, August 1). Designing Efficient User Interfaces For Games. Medium. https://medium.com/@nicolaskraj/designing-efficient-user-interfaces-for-games-be20b516f1c2

The research of this thesis will focus on the diegetic and non-diegetic markers of the tangible UI of the VR games. This constrain is introduced due to the limited timeframe of this research and lack of funding. This is an opportunity for the future researchers to expand this topic and focus on other types of UI. Additionally, diegetic and non-diegetic represent the two opposites and testing on them helps to establish whether there is a relationship between a marker and the narration immersion at all.

Non-diegetic UIs are not the part of the game world, and they are not visible for the avatar of the game. They are displayed only for the player, usually as a graphic overlay on the screen. They are often used to display information that only the payer needs to know e.g. inventory, health or amount of in-game currency.

Diegetic UI looks like it's part of the virtual world and thusly makes the play more immersive (Andersson, 2024; Fagerholt, Lorentzon, 2009). An example of diegetic UI in use could be displaying information in a game-world object, like presenting the health bar of the main character in Dead Space 2 on a mechanic spine that he is wearing.

While the diegetic and non-diegetic framework provides a valuable lens for categorising UI elements based on their relationship to the game world, it does not fully explain *how* players recognise that they can be interacted with in the first place. Such recognition relates to the theory of affordance that will be explained in detail in the 2.2.1 section.

2.1.1 Diegetic vs Non-diegetic UI elements for VR experiences

In the past there has been a few attempts to test how different UI design influences players in the VR games following the Fagerholt and Lorentzon's theory. Multiple researchers have studied diegetic vs non-diegetic UI in VR, but they all focused on shooters and spatial

presence or user preference rather than the narration immersion. Nonetheless they are worth acknowledging and analysing in relation to this research.

Saling et al. (2021) and Köhle et al. (2021) both examined diegetic versus non-diegetic UI in VR shooter games - the former measuring spatial presence, the latter measuring user preference. Neither study found significant effects of UI type on their measured outcomes: Saling et al. could not establish a strong relationship between UI design and sense of presence, while Köhle et al. concluded that different UI approaches were simply suited to different contexts. Both studies also faced methodological constraints, with sample sizes of 13 and 37 participants respectively.

These findings raise an important question for the present research. If UI design type does not strongly affect spatial presence - the feeling of physically "being there" - does it follow that narrative immersion is equally unaffected? There is reason to hypothesize otherwise. Spatial presence in VR is largely achieved through hardware (the headset's visual and auditory immersion), making it potentially robust to interface variations. Narrative immersion, by contrast, requires sustained cognitive and emotional engagement with a story - a state that may be more vulnerable to disruption when UI elements break the fourth wall by reminding players they are in a game. The immersion types will be explained in detail in the section 2.3 of this thesis. Furthermore, both studies tested action-oriented shooter games where narrative engagement was minimal. How UI design affects players in story-driven VR experiences remains empirically untested.

To conclude, while these studies contribute valuable insights into UI design preferences and spatial presence in VR shooters, their validity limitations and narrow focus leave unaddressed the crucial question of how UI design choices influence narrative immersion in story-driven VR experiences, particularly when considering the affordances of tangible interface markers.

2.2 RECOGNISING INTERACTABLE OBJECTS IN GAMES

User interface consists of large range of elements such as menu, health bar or buttons. As explained in the Introduction chapter - in the VR environments UI can be symbolic or textual, and the focus of this research is on the tangible spatial UI, such as the 3D objects in the virtual environment. The researcher of this thesis decided to focus on the "markers" of interactable objects in games that let players know if they can interact with one object but not another. The markers used to indicate interactable objects in VR environments are essentially tools for communicating affordances to players, bridging the gap between the object's inherent interactive capabilities and the player's perception of those capabilities.

2.2.1 Affordance and Semiotics

The ecological psychologist James Gibson, who originated affordance theory (1966 and 1979), characterized affordances as the "opportunities and possibilities that an environment offers to individuals"[9]. This concept emphasizes how the physical characteristics of objects in

[9] Wikipedia contributors. (2025, October 9). *Affordance*. Wikipedia. https://en.wikipedia.org/wiki/Affordance

relation to a user's capabilities reveal potential interactions through observable properties. For example, an object positioned at knee level with a rigid, flat, and sufficiently broad top surface possesses characteristics that suggest the possibility of sitting. The user perceives and behaves by observing the information of the subject without previous knowledge or instructions (J. Lee et al., 2018). Additionally, Jakob Steffen et al.'s "Framework of Affordances for Virtual Reality and Augmented Reality" identified fundamental affordances inherent to the VR medium, validating their theoretical framework through empirical studies.

On the other hand, Norman (1988) approached affordances from a user-centered perspective rather than viewing them as objective environmental features, conceptualizing them as both the "apparent and actual properties that objects present to users"[10]. This user-focused interpretation became foundational for the Human-Computer Interaction (HCI) field and broader mediated communication, establishing affordances as a central concept for understanding how people interact with digital systems.

Norman differentiated between the terms that he introduced - "perceived affordances" and Gibson's concept of "actual affordances". The interpretation of the object from each user's experience does not necessarily match the intended property or design of the object. A button may afford clicking, but if users do not perceive it as clickable, the affordance goes unused. Norman later refined this further by introducing the term *signifier* to describe design elements that communicate affordances to users (Norman, 2013). While an affordance is the possibility for action, a signifier is what makes that possibility visible.

The concept of "signifiers" originates in Ferdinand de Saussure's semiotics, where a signifier is the perceivable form (a word, image, or symbol) that communicates meaning to an interpreter. Norman adapted this concept for design contexts, using "signifier" to describe elements that communicate affordances to users. However, not all signifiers work the same way. Semiotic theory distinguishes between *conventional signs*, which require learned associations (such as an exclamation mark icon that gamers recognize as "interact here"), and *motivated signs*, which communicate through natural perceptual properties (such as brightness or contrast that instinctively draws the eye). This distinction is directly relevant to marker design in VR: a non-diegetic marker like a floating icon relies on players' familiarity with gaming conventions, while a diegetic marker like a subtle glow leverages perceptual principles that function independently of prior gaming experience. Both signify the same affordance — that an object can be interacted with — but through different semiotic mechanisms that may affect narrative immersion differently.

This distinction between affordances and signifiers maps directly onto the relationship this thesis investigates. In a VR environment, interactable objects possess affordances - a toy can be picked up. However, players can only act on these affordances if they perceive them. Markers, in the terminology of this thesis, function as signifiers: they are the visual design elements that communicate to players which objects possess interactive affordances. The

---

[10] *What are Affordances?* (2025, September 19). The Interaction Design Foundation. https://www.interaction-design.org/literature/topics/affordances

research question thus becomes not whether affordances exist in the virtual environment, but how effectively different marker designs signify those affordances while maintaining narrative immersion.

The evolution of the affordance theory, from Gibson's ecological psychology to Norman's user-centered design perspective provides a rich framework for understanding player interaction in VR environments. Affordance theory offers a valuable perspective for examining how virtual objects have similar affordances to their physical word counterparts. For the research methodology chapter, the researcher of this thesis will use the VR experience that they've made in the past titled "Echoes of Abandonment". The game was explained in the introduction, and interactable objects in this space were toys. In the context of this research paper, affordance is the fundamental property that the object *can be picked up* in the VR environment. Knowing affordance theory, the researcher is investigating how different marker types affect the perception of affordances, which in turn affects narrative immersion.

2.2.2 Markers

Certain ways to design the UI in a game, following the diegetic theory, can signal the affordances to the players more or less effectively. Markers are the UI design elements the designer is adding to communicate or signal that an affordance exists. They are used to "mark" with which objects in the virtual 3D space we can interact and in what way. For example, the affordance is "pick-up ability" of a toy in a VR environment that simulates a playroom. Markers can be used to ensure that the player knows what objects they can pick up. The design can be diegetic – for example a subtle glow emanating from the object, particles floating around it, or the object appearing slightly different in texture or colour within the game world; or non-diegetic – an icon hovering above the object, an outline shader applied to it, or a UI prompt saying "Press X to pick up". Figure 2.3 below shows an example of what a marker can look like in the VR game "Knight in the attic" by Mighty Yell (2023).

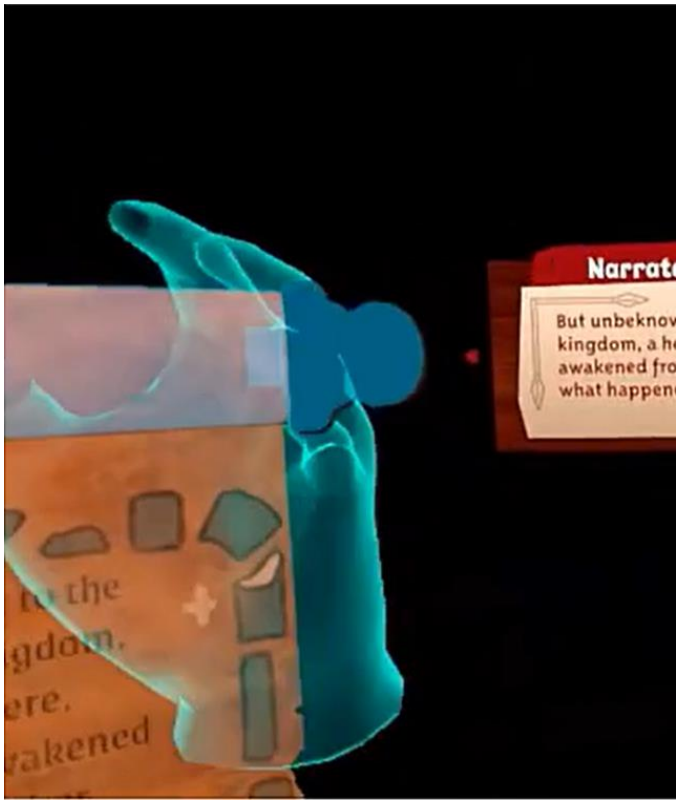

Figure 2.3 - Highlighting interactable part of an object with a virtual hand on hover in VR. Source: Knight in the Attic by Mighty Yell, 2023.

Makers function as the designed mechanism to bridge the gap between actual affordances (what is pragmatically possible) and perceived affordances (what players believe they can do). The critical question of this paper is not whether the affordance exists, but how the visual communication of that affordance - through diegetic markers integrated into the story world, non-diegetic UI overlays, or no explicit markers at all - influences players' ability to remain transported into the narrative while simultaneously recognizing interaction opportunities.

2.2.3 Designing for Usability

Keeping the user experience (UX) in mind, the researcher of this thesis should use the theories discussed in the literature review chapter to ensure satisfactory experience of the VR game to the participants of my study. Preece et al. (2002) explain that user experience occurs as a result of achieving usability goals during an interaction. System usability refers to the interface design of a given application (Roth & Koenitz, 2016). In the context of this research non-diegetic markers may excel at usability goals by clearly communicating affordances with high effectiveness and efficiency, while potentially compromising user experience goals related to immersion and emotional fulfilment by breaking the narrative engagement. On the other hand, diegetic markers may preserve experiential qualities but risk reducing usability through decreased clarity of affordance communication. Understanding this balance is critical for VR user interface design.

According to Preece, Rogers, and Sharp's framework (2002), usability goals are at the center of Interaction Design, while user experience goals form the outer ring of the diagram presented on the figure 2.4.

Figure 2.4 - Diagram on usability goals and user experience by Jenifer Preece. Source: Raffaele, R., Carvalho, B., Lins, A., Marques, L., & Soares, M. M. (2016). Digital Game for Teaching and Learning: An analysis of usability and Experience of Educational games. In Lecture notes in computer science (pp. 303–310). https://doi.org/10.1007/978-3-319-40355-7_29

This framework provides the theoretical support that would guide the researcher in the development of markers that will be discussed in the 3.1.2.1 section. Preece et al.'s framework reveals that effective interface design must satisfy both usability goals and user experience goals. In VR narrative games, user experience goals such as emotional fulfilment and enjoyment are closely tied to narrative immersion - the very construct this thesis investigates. By testing both marker types alongside a control condition, this research can empirically examine whether the theoretical trade-off between usability and experience manifests in measurable differences in narrative immersion.

2.2.4 Visual Interaction Cue Framework in the games industry

While affordance theory establishes what interaction possibilities exist and why users perceive them, a crucial gap remains in understanding how games visually communicate these affordances to players. Dillman et al. (2018) address this gap by systematically analysing visual interaction cues across 49 video games, introducing a classification that distinguishes cues along two key dimensions: their integration with the game environment and their obtrusiveness to the player.

Their framework identifies "subtle" cues - those using lighting and environmental contrast to guide attention while remaining world-integrated - as distinct from "emphasized" cues that explicitly highlight objects through overlays, outlines, or symbolic markers. This empirical classification parallels the diegetic/non-diegetic framework established by Fagerholt and Lorentzon (2009), but grounds it in systematic observation of actual game design practices rather than theoretical categorization alone. The marker designs selected for the present study align with Dillman et al.'s empirically-derived categories: the brightness shader functions as a subtle, world-integrated cue, while the exclamation mark icon operates as an emphasized, obtrusive cue.

Critically, Dillman et al.'s research focused on classifying design patterns rather than measuring their effects on player experience. Their work, conducted in AR contexts, does not examine how these different visual communication strategies affect narrative immersion in VR storytelling environments. However, combining their empirical classification with the

theoretical frameworks established in this thesis — particularly the distinction between motivated and conventional signs (Norman, 2013; Saussure) and the vulnerability of narrative transportation to disruption (Green & Brock, 2000) - generates a testable prediction. Subtle, world-integrated cues that function as motivated signs should preserve narrative immersion by maintaining the coherence of the story world, whereas emphasized, obtrusive cues that function as conventional signs may disrupt immersion by reminding players they are interacting with a game interface. This prediction forms the theoretical basis for Hypothesis 1 of the present study.

2.3 IMMERSION

The ability to stimulate "immersion" in video games is one of the most noted characteristics of this medium. There is no one concrete shared definition stating what immersion is (Wilkinson, Brantley, & Feng, 2021). Jennett et al. (2008) proposed that one experiences immersion as "a loss of awareness of time, real world involvement, and sense of being in the task as a consequence of a good "gaming experience"". In the current research field, there are a few types of immersion categorised. Haggis-Burridge (2020) proposes the following distinction.

SYSTEMS IMMERSION	SPATIAL IMMERSION	EMPATHIC/SOCIAL IMMERSION	NARRATIVE/SEQUENTIAL IMMERSION
A high level of engagement with the systems and decision-making processes in the game, related to 'flow'.	A sense of 'presence' in a location. The feeling of being in that place, or of having travelled there.	An emotional connection with the characters or social context of a game.	A deep and compelling investment in the progression of events, locations, and/or abilities. The focus here will typically be 'what happens next?'
All types are likely to be non-discreet, with close relationships and overlaps of game-elements that contribute to (or subtract from) multiple forms of immersion.			

TABLE 1: THE FOUR CATEGORIES OF IMMERSION.

Figure 2.3 – The Four Categories of Immersion. Source: Haggis-Burridge, M. (2020). Four categories for meaningful discussion of immersion in video games. Pure Buas.
https://pure.buas.nl/ws/files/1228916/Haggis_Immersion_4CategoriesInVideoGames.pdf

Immersion is not a static experience but describes a scale of involvement with a game (Brown & Cairns, 2004). Those researchers introduced the term "total immersion" that is described as "a Zen-like state where your hands just seem to know what to do, and your mind just seems to carry on with the story."

Following the distinction presented in the Figure 2.3, the researcher of this thesis decided to focus on the narrative immersion. One reason for this choice is that historically, VR research has over-emphasised technological variables and spatial presence (Pianzola, Riva, Kukkonen and Mantovani (2021); Hvass, Larsen, Vendelbo, Nilsson, Nordahl and Serafin (2017)). Immersion through story engagement requires sustained cognitive and emotional engagement vulnerable to interface design choices. Lastly, as mentioned in the introduction, this choice

comes from personal interest of the researcher as proven by their bachelor thesis "Storytelling in Interactive Environments".

A reason to investigate narrative immersion is because storytelling represents a core aspect of human life. Different mediums—including cinema, journalism, and advertising, or even sports—continuously present narratives to make consumers engage more with a product (Júnior et al, 2022). Yet stories also operate at an intimate level: some of our stories are very personal: "the story of our life" isn't just a metaphor, but something we actually tell ourselves and other people to give shape and meaning to who we are (Glassner, 2004).

2.3.1 Narration immersion

This thesis will be focusing on the immersion to the narration. Scientific interest in narrative processing and its effects has substantially increased in recent years, with narratives recognized as powerful means of engaging audiences across multiple media forms (Green and Appel, 2024). With scholars using terms such as presence, flow, absorption, and engagement interchangeably or in overlapping ways (Nilsson, Nordahl & Serafin, 2016) it becomes challenging to investigate how UI design affects player experience: if immersion encompasses everything from technological sensory input to psychological story engagement, it becomes difficult to isolate what specific design choices actually influence. For the purposes of this research, I distinguish between these related concepts to arrive at a precise operationalization of narrative immersion.

2.3.1.1 Transportation Theory Model

Narrative transportation model (Green and Brock, 2000) describes transportation as an experimental state in which one's mental processes (attention, emotion, and imagery) are concentrated on the events occurring in the narrative. Transported individuals mentally leave the real world behind and become entirely focused on the story world[11]. The transportation scale assesses dimensions of transporting, including emotional involvement, cognitive attention, suspense, lack of awareness of surroundings and mental imagery.

Admirably, this theory has been developed and tested on readers with books, rather than in an interactive narration of a VR game. Therefore, the next step is to look for research that focuses on games, or games related medium.

2.3.1.2 Flow and Cognitive Absorption

Flow theory (Csikszentmihalyi, 1975) and cognitive absorption (Agarwal & Karahanna, 2000) are sometimes used interchangeably with the concept of immersion, but describe distinct psychological states. Flow refers to intense focus achieved when task difficulty balances with skill level - a state applicable to any activity, not specifically narrative engagement. Additionally, the root of cognitive absorption lays in the flow theory by Csikszentmihalyi, but unlike it, it is not referring to the state of intense concentration, but rather deep involvement within the

[11] Wikipedia contributors. (2025b, October 8). *Transportation theory (psychology)*. Wikipedia. https://en.wikipedia.org/wiki/Transportation_theory_(psychology)

technology interaction, where cognitive absorption leads to user experience rather than being the same thing as user experience.

Procci and Bowers (2011) empirically demonstrated that immersion and flow function as unique constructs, despite surface similarities. By using two measuring scales that are associated with those concepts – Immersive Tendencies Questionnaire (ITQ) and Dispositional Flow State Scale (DFS-2) - they found that even though those questionnaires had some similarly phrased questions, they gave results that proved immersion and flow being unique constructs. Neither concept captures what this thesis investigates: sustained engagement with narrative content in an interactive VR environment.

2.3.1.3 How to measure narration immersion in VR?

Building on transportation theory, Busselle and Bilandzic (2009) developed the Narrative Engagement Scale, which measures dimensions including narrative understanding, attentional focus, emotional engagement, and narrative presence. This scale offers a more applicable framework for interactive contexts than transportation scale that was designed for reading. However, narrative engagement alone does not account for the specific qualities of VR as a medium.

The Immersion Experience Questionnaire (Jennett et al., 2008; Cutting et al., 2025) addresses this gap by measuring immersion specifically within game contexts, assessing factors such as cognitive involvement, emotional involvement, and real-world dissociation. While not designed specifically for narrative, this instrument captures how players experience immersion in interactive digital environments, including VR.

For this thesis, narrative immersion is operationalized as the combination of narrative engagement (investment in story content) and immersion experience (absorption in the VR game context). Neither construct alone fully captures what it means to be immersed in a VR narrative: narrative engagement measures story involvement but was not developed for VR, while immersion experience measures VR game absorption but is not narrative-specific. Together, they provide a comprehensive measure of the dependent variable this research investigates - a player's sustained cognitive and emotional engagement with a story told through an interactive VR environment.

2.3.2 Narrative immersion research for VR

In the context of VR immersion, narrative immersion hasn't been studied as much as other immersion types. Brooks (2003) argues that immersive environments existed long before computers through oral storytelling. In a virtual reality environment that uses narrative, the participant is doubly immersed - both sensually (surrounded by computer-controlled media) and narratively (engaged in the story world). This "double immersion" makes VR a powerful medium for testing narrative immersion of players.

Pillai and Verma (2019) have conducted research on narrative immersion in VR cinema, what doesn't have the same interactivity level as VR games that are the focus of this thesis, but nonetheless present interesting results. They distinguish between technical immersion (achieved through attentional cues and environmental acknowledgment) and narrative immersion (arising from story, character, and viewer integration). Essentially, they found that technical and narrative aspects can support each other to strengthen overall immersion,

suggesting that UI markers need not necessarily disrupt narrative engagement if properly integrated within the story world.

2.3.3 Interactive Narrative in VR

Abbott (2002) defines narrative as "a representation of a story, involving a sequence of events and characters that unfold over time". This definition was written for traditional media, such as film, theatre, or literature. Video games, or VR games have a major characteristic that make the narrative experience unique.

Narrative immersion comes from player's eagerness to see the progression of a story that they are experiencing the game. Typically, it resolves around sequence of events in a traditional story, but it can be also applied to progressive upgrading of a character or unlocking new regions in a game to explore (Haggins-Burridge, 2020). As mentioned before, game stories have their own unique characteristics. One of the most important characteristics of VR narrative is that it offers its players more interactable experience than other popular mediums[12].

Interactive Digital Narratives (IDNs) are defined as "a form of narrative expression in the digital interactive medium" (Roth & Koenitz, 2016), distinguished from traditional media by their interactive nature and digital delivery. Roth & Koenitz mention that the IND puts agency on interaction and the ability to influence important aspects of a narrative (like character development, sequencing, or outcome), allowing players to participate in the creation of their own narrative experience mediated by a computer system.

The interactability of a VR narrative experience is not only referred to by the actions of the player, but also a passive experience of it. In both 6DoF (six degrees of freedom) experience and a VR360 Video, players have the power to control the camera in any angle they would like what already gives them more control than in some games or in movies (Dooley, 2017). The sole act of wearing a headset on which a player can experience the story is already giving them more agency.

Brooks (2003) in their research points out that it is unnecessary to design complex visual effects or scenery for a particular part of a VR experience if the audience is already deeply engaged in narrative events.

In 2020 two scientist from Science University of Norway – Irshad and Perkins – attempted to study how to increase user engagement in VR games by introducing their players to interactive narratives. They conducted a study on 32 participants where they were asked to perform tasks in two immersive environments. One of them being a 3D landscape with no narrative story connected to the experience, second being the same landscape with narrative tasks that need to be performed by the participants. "Participants in Group1 were asked to explore and interact with the environment; however, no particular tasks were assigned (...) participants in Group 2 were asked to perform a set of tasks i.e., find an underground tunnel and remove the obstructions to enter the tunnel to save themselves from the flood." The independent variables of the study were narrative and non-narrative VR immersive

[12] Qin, H., Patrick Rau, P. L., & Salvendy, G. (2009). Measuring Player Immersion in the Computer Game Narrative. *International Journal of Human–Computer Interaction*, *25*(2), 107–133. https://doi.org/10.1080/10447310802546732

environments. The level of engagement inside the VR was used as the dependent variable. Results showed a significant increase in the overall level of engagement in a narrative-based VR environment.

This study shows narrative does increase engagement - which supports part of my thesis premise; however, it didn't examine how UI design or markers might affect that narrative engagement. Ultimately, as this research is relevant for the field of narrative immersion, it focused on a unique variable of user engagement rather than immersion itself. Due to missing of the key components that I would like to investigate in this thesis – the influence of markers design and low statistical significance, I am highlighting the research gap of my chosen field.

2.4 SUMMARY

The research discussed so far presents a visible gap in the field of investigating narration immersion in the VR games and the influence of markers design on maintaining or breaking it. Therefore, this thesis will answer the question – what the influence of markers design on the player's narration immersion in VR games is. The answer to this research question could be useful for the industry experts that design VR environments or UI on daily basis, researchers who look for answers to this research gap, and personally I find it interesting as an aspiring game designer.

2.4.1 Ethical considerations

One of the focus points of this research is immersion, that is a concept that carries some ethical concerns with it. Although immersion can enhance how players engage with games, it has been linked to harmful social consequences (Seah & Cairns, 2008). The example of an extreme case involved parents so absorbed in an online game that they neglected their infant, resulting in the child's death from starvation[13]. This drastic incident likely stems from addiction rather than immersion itself, although immersive experiences can contribute to gaming addiction. Video game addiction is a condition characterized by severely reduced control over gaming habits, resulting in negative consequences in many aspects of your life, including self-care, relationships, school and work. The aim of this study is not to contribute to any addictive behaviour of its players, rather measure the narration immersion to help with better storytelling in interactive environments.

While choosing VR as a medium for research, one must acknowledge some limitations that it inherits. Virtual Reality systems depend heavily on sensory immersion, engaging multiple sensory channels simultaneously (Berkman & Akan, 2019), and typically assume users possess full body mobility. Furthermore, VR has been causing motion sickness to many users and research that investigates it is still ongoing (Chattha et al, 2020). The findings regarding UI marker design and narrative immersion may therefore primarily benefit users without these physical limitations. Future research should investigate how different affordance signalling approaches affect narrative engagement for users employing alternative interaction modalities, such as seated experiences, reduced movement interfaces, or adaptive control schemes.

[13] Bavelier, D., Green, C. S., Han, D. H., Renshaw, P. F., Merzenich, M. M., & Gentile, D. A. (2011). Brains on video games. Nature reviews. Neuroscience, 12(12), 763–768. https://doi.org/10.1038/nrn3135

As mentioned, the research will be conducted on a one-woman game made for the bachelor graduation project by the author of this thesis. Main reasons for using the "Echoes of Abandonment" for the methodology is accessibility to the pre-build and resealed game, and full control over the virtual world that could prevent any unwanted for the research factors, such as distracting stimuli. Story told in this game however might be triggering to some participants, therefore it is important to take their mental health into consideration. Potential triggers could be racism, homophobia, transphobia, (sexual) violence, (child) abuse, self-harm, suicide, animal abuse, eating disorders and sizeism or pregnancy related issues (Healthline, 2019). The story of this VR experience is focusing on a child experiencing their parents' fighting, so it is important to remember to inform the participants of the study sufficiently before they take part in the experiment.

2.4.2 Keeping in mind

In her book "Hamlet on the Holodeck: The Future of Narrative in Cyberspace" (1997), Janet H. Murray examines immersive experiences, which she describes as fragile and easily disrupted[14]. This fragility becomes particularly relevant when examining UI design choices in VR: markers that draw attention to game mechanics risk breaking the immersive spell that narrative engagement requires. She says "Giving the audience access to the raw materials of creation runs the risk of undermining the narrative experience.". "Murray's (1997) observation that exposing the 'raw materials of creation' risks undermining narrative experience applies directly to non-diegetic UI markers. Following this reasoning, the researcher of this thesis decided that to utilise a "non-disturbing" markers design for both diegetic and non-diegetic test group. This design choice will be described in detail in the methodology chapter.

2.4.3 Hypothesis

To conclude all the literature review presented in this thesis the following hypothesis will be posed to prove with the methodology described in the chapter 3:

H1 The diegetic markers will support the narration immersion.

H2 The non-diegetic markers will noticeably disturb the narration immersion of the VR game players.

Empirical research supports this hypothesis, though with important limitations. Iacovides et al. (2015) found that removing non-diegetic HUD elements increased immersion for expert players in traditional video games. Salomoni et al. (2016) demonstrated that diegetic interfaces in VR (using Oculus Rift) enhanced players' feeling of "being in the flow of the game." However, neither study measured narrative immersion specifically — Iacovides et al. examined player involvement in non-VR contexts, while Salomoni et al. assessed presence rather than story engagement. Whether these findings extend to narrative immersion in contemporary VR environments remains untested.

[14] Wikipedia contributors. (2025a, July 13). *Hamlet on the Holodeck*. Wikipedia. https://en.wikipedia.org/wiki/Hamlet_on_the_Holodeck

Based on this theoretical and empirical foundation, this thesis hypothesizes that diegetic markers will result in higher narrative immersion than non-diegetic markers in VR game environments. The following methodology chapter describes how this hypothesis will be tested.

3 METHODOLOGY

The research of this thesis is concerned with the influence of the markers design on the narrative immersion of the VR games players. This research will be based on a questionnaire, since those are most used in VR evaluation due to their ease-of-use and versatility in topics (Halbig & Latoschik, 2021) and enable evaluation of subjective lived experiences (Grassini & Laumann, 2020).

To investigate the hypothesis 1 and 2, this study has been designed in the following manner. The participants will be randomly assigned into 3 conditions. All of them will be experiencing the same VR game "Echoes of Abandonment", with different marker designs. Group A will be presented an experience with non-diegetic markers, Group B will have diegetic markers, and Group C will be the control group with no markers on the interactable objects. After experiencing the game, each participant will be asked to fill out a questionnaire designed to help answer the topic of this research. The whole methodology in detail will be explained in the following sections.

3.1 METHODOLOGY CHOICE

This study will be using a questionnaire as its main methodology. This choice is partially motivated by its popularity among other VR immersion research, which will ensure the data gathered is comparable to already existing research. Researchers frequently use questionnaires to evaluate the User Experience (UX) mainly due to their cost-effectiveness, systematic nature, and ease of application (Bareišytė at al., 2024). Additionally, concepts that the research question of this paper is concerned with, are measured with the questionnaires (immersion and narration engagement).

Alternative methodologies that could have been used in this research are Heuristic evaluation, A/B testing or Usability testing. According to "*Universal methods of design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*" by Martin & Hanington (2012), those are the definitions of those methodologies: "A/B testing is an optimization technique that allows you to compare two different versions of a design to see which one gets you closer to a business objective. (...); A heuristic evaluation is an informal usability inspection method that asks evaluators to assess an interface against a set of agreed-upon best practices, or usability "rules of thumb."; The goal of a [usability] report, (...) is to clearly outline which parts of the user interface should be fixed or improved".

The nature of the research question investigates narration immersion - a subjective psychological state that cannot be directly observed. A/B testing, which compares two design versions to determine which better achieves a specific objective (Martin & Hanington, 2012), does not fit this study for two reasons: first, this research compares three conditions rather than two; second, the goal is not to identify which marker type players prefer, but to understand how each marker type influences narrative immersion.

Heuristic evaluation, which asks experts to assess an interface against established usability best practices (Martin & Hanington, 2012), was also unsuitable because this research does not evaluate whether markers conform to design standards - rather, it examines how different marker designs affect players' subjective experience.

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

Similarly, usability testing focuses on identifying interface problems that need improvement (Martin & Hanington, 2012), whereas this study investigates the relationship between design choices and psychological outcomes, not interface deficiencies.

Questionnaires, by contrast, allow participants to report their own subjective experiences of immersion and narrative engagement - the precise constructs this research aims to measure.

3.2 MATERIALS

3.2.1 Questionnaire Development

To measure the narration engagement the Narration Engagement Scale (Busselle and Bilandzic, 2009) will be used, and for the immersion of the players, the Immersion Experience Questionnaire (Jennett et al., 2008) has been implemented. Keeping the time and comfort of the participants of this study in mind, it has been decided to use the shortened version of the IEQ, since it's been academically validated (Cutting et al., 2025).

Questions from the Immersive Experience Questionnaire Short Form (IEQ-SF) focus on one's experiences within a game (Jennett et al., 2008; Cutting et al., 2025) rather than general experiences and abilities to feel involved or switch between tasks, as in the ITQ (Witmer & Singer, 1998). IEQ-SF measures three distinct dimensions of game immersion: Involvement, Real-World Dissociation, and Challenge. This three-dimensional structure is useful to this research because marker design may differentially influence these aspects of immersion; for instance, markers that are too visually intrusive might disrupt Real-World Dissociation by constantly reminding players they are in a game interface, while markers that fail to communicate affordances clearly might negatively impact the sense of Involvement by creating confusion or frustration.

The Short version of the IEQ has been chosen for this study to ensure comfort of participants having to fill out an 11-item questionnaire rather than 31, therefore avoiding the survey fatigue (O'Brien, 2016). Moreover, Cutting et al. (2025) themselves motivated the development of the IEQ-SF by noting that the original IEQ (Jennett et al., 2008) had been developed over a decade earlier using methods that may no longer meet current standards of transparency and validity. These issues suggest that the IEQ, which was developed using older analysis and data gathering tech-niques may not live up to modern standards of transparency and validity. To quote the Cutting et. Al (2025) "The IEQ-SF is thus suitable for use in all cases where the IEQ was previously used, but is particularly suitable for use in studies which need participants to fill in several scales, where the reduced size of the IEQ-SF could help prevent the study becoming too long and participants from suffering questionnaire fatigue."

The Narrative Engagement Scale (NES) (Busselle and Bilandzic, 2009) measures four distinct dimensions of experiential engagement with narratives: Narrative understanding, Attentional focus, Emotional engagement, and Narrative presence. This multidimensional approach is particularly suited to the present research because the marker design conditions may differentially affect these dimensions; for instance, non-diegetic markers might disrupt narrative presence while leaving narrative understanding intact, whereas diegetic markers may preserve the sense of being "inside" the story world.

In addition to those academically validated scales, the researcher of this study added her custom questions to help with the answering of the hypothesis of this study. Those questions are listed below, and full questionnaire can be found in Appendix B.

1. *What condition were you part of?*
   -diegetic
   -non-diegetic
   -control group

2. *It was intuitive to understand how I could interact with objects in the virtual environment*
   strongly disagree – strongly agree

3. *Sounds that played when I interacted with the toys enhanced my enjoyment of the game*
   strongly disagree – strongly agree

4. *The marker of interactable toys made it easy for me to recognize which toys I can pick up*
   strongly disagree – strongly agree

5. *I found it easy to follow the story while playing the game*
   strongly disagree – strongly agree

6. *The interaction markers helped me understand what was important to the story.*
   strongly disagree – strongly agree

7. What is your level of experience with VR games?

8. *What is your profession / specialty?*

9. *What did you like about the design of the markers present in your VR experience?*
   -open question

The total questionnaire consists of 32 questions – 11 questions from the IEQ-SF, 12 questions from NES and 9 custom questions from the researcher. Using academically validated scales ensures measurement reliability and allows comparison with existing literature. All statement-based questions (1-30) use a 7-point Likert scale from "Strongly Agree" to "Strongly Disagree" to maintain consistency across the questionnaire and reduce cognitive load for participants. The custom questions introduced by the researcher are based on the literature review and are meant to support proving of the Hypothesis 1 and Hypothesis 2.

Question 1 will help the researcher in the data analysis by establishing clearly what the condition of the participant was taking part in this study.

Question 2 serves to introduce a control variable to verify baseline usability across conditions. The question measures how effectively the virtual environment communicates interaction possibilities. According to Norman's (1988) concept of perceived affordances, the question essentially measures whether the virtual environment's affordances were successfully perceived regardless of marker condition. The researcher expects no significant difference

between conditions. Therefore, if no difference is found, it confirms marker designs didn't create usability concerns, validating that any narrative immersion differences are due to marker type rather than poor interaction design. However, if unexpected differences emerge, then it reveals potential concerns that must be discussed in limitations.

Question 3 verifies that audio design (a core game mechanic) didn't distract from narrative immersion. In "Echoes of Abandonment," players hear parents arguing behind the door; each toy creates musical sounds that compose together to mute the argument making the sound narratively and mechanically central. While sound is not an independent variable being manipulated, it's crucial to verify that it functioned as intended. According to Murray (1997), immersion is sustained through multiple sensory channels, therefore researcher expects no influence of sound on immersion.

Question 4 This question directly operationalises the effectiveness of the markers as signifiers, what Norman (1988) defines as perceivable indicators that communicate the appropriate interaction possibilities of an object. It measures whether each marker type successfully fulfilled its intended function of communicating affordances in the virtual environment, a construct that Steffen et al. (2019) validated as central to user interaction in VR contexts. Differences between conditions on this item would indicate that the markers varied in their communicative clarity, which is essential context for interpreting the primary immersion findings.

Question 5 measures cognitive load related to narrative processing. One might notice the similarity to the question 2 from the Narrative Engagement Scale that states "I had a hard time recognizing the thread of the story". Both questions measure related but distinct constructs - one captures whether participants understood the story (outcome), the other captures how effortful it was to maintain that understanding (process). This distinction is grounded in Green and Brock's (2000) transportation theory, which posits that cognitive effort required to follow a narrative can disrupt the state of narrative transportation, even when comprehension is ultimately achieved. If marker type influences the ease of narrative processing without affecting comprehension outcomes, this item would capture that difference where the NES alone might not.

Question 6 is a core question of this questionnaire. It directly tests whether marker design influences narrative immersion and captures the link between UI design choices and story understanding. It is grounded in Fagerholt and Lorentzon's (2009) diegetic/non-diegetic framework: diegetic markers, by existing within the story world's logic, should theoretically feel integrated with narrative meaning, while non-diegetic markers impose a layer of symbolic mediation (Saussure, 1917) that may either clarify or disrupt the relationship between interaction and story.

Question 7 is present in the questionnaire because the familiarity with the VR technology might influence how immersed people can be in the story rather than confused about the functionality of the headset.

Question 8 will help in testing any influence of player's profession on their ability to recognize UI elements in the VR game. As much as this concept is not part of the research question, the researcher is aware of the possible bias that comes from the sampling method used in this research.

Question 9 is an open question that may reveal unexpected insights about marker perception. Qualitative responses can help explain patterns in quantitative data or reveal dimensions not captured by structured questions. Qualitative data complements quantitative findings by providing contextual information that helps explain why certain patterns emerged or failed to emerge (Lankoski & Björk, 2015). These responses are analysed thematically in Chapter 4.

3.2.2 Game Development

The study was conducted on the game made by the researcher of this thesis "Echoes of Abandonment", that has been described in the Introduction chapter. This VR experience is built in Unity version 2022.3.22f1 and runs on the ASUS gaming laptop with the graphical card RTX 3060 and 11th Gen Intel(R) Core(TM) i7-11370H. Reason for choosing this type of equipment is to ensure the high quality of performance of the game, since possible lagging or low resolution could influence the VR motion sickness (Kundu et al., 2021). For the same reason, the VR headset used in all study conditions was Meta Oculus 3, due to the accessibility of this hardware to the researcher, and providing good quality of VR experience for the participants (4K image).

The virtual environment consisted of handmade, clay-sculpted kid's bedroom, scanned used photogrammetry, and uploaded to Unity as a virtual environment. In the virtual room there were placed interactable toys that are the focus of this research. The interactable objects in were toys from open source free to download online database. Toys included:

1. Plushie Elephant
2. Plushie Lion
3. Plushie Duck
4. Plushie Dino
5. Wooden Plane
6. 6 Crayons
7. 5 wooden letter blocks spelling "F-A-M-I-L-Y"
8. Space Rocket
9. Kazoo
10. Xylophone and a baton to play it

Listed 3D objects has been selected to be the tangible UI element of this VR experience because of their compliance with the story world. Criteria for the selection of what objects would be implemented follow the theoretical background presented in the literature review. Following the affordance theory the 3D objects needed to signal the intended interaction in a way that can be interpreted by the players as designed by the researcher. The selected toys in the real world are meant to be picked up to play with, and as mentioned in the introduction the interaction that this research is investigating is based on the pick-up interaction. Additionally, following the semiotics and usability guidelines the tangible UI needed to be easily interpreted by the players to ensure their easy interpretation of the virtual world. Lastly, those particular toys have been selected due to their accessibility in the free database of 3D toys at the time of making of the game.

Part of the virtual environment story was hearing parents' argument behind the door. The full script has been written by the researcher and can be found in the Appendix A. The parents were played by two voice actors that were selected due to them being native English speakers, and their relationship with the researcher. Participants, upon picking up a toy, would

hear a sound play that is attached to the toy. Interacting with toys would create musical composition that would quiet down the fight of parents. Music and sounds were in majority created by a composer Jasmine Karimova.

The game "Echoes of Abandonment" has been displayed as part of the graduation show of the Royal Academy of Art in the Hague of 2023/2024 graduating students. During the 4 days of exhibition, it is estimated that ~100 people have experienced this game, aged approximately between 10 and 80. Following photographs showcase the virtual environment and external presentation.

Figure 3.1 What the players see in VR with no markers condition - Screenshot from inside the game "Echoes of Abandonment" by Paulina Cywoniuk, 2024

Figure 3.2 Player of the VR game during the exhibition in the Royal Academy of Art, the Hague - "Echoes of Abandonment" by Paulina Cywoniuk, 2024

The exhibition consisted of physical space of a circle (blue carpet visible on the Figure 3.2) of the measurements 2x2 meters. This physical constrain has been implemented to ensure

players' safety in the exhibition space, and the safety of the artworks presented in the same room. In the current study similar constrains will be implemented. Instead of using 2x2 carpet in the actual study the researcher will be using the virtual boarder that the Oculus 3 has embedded in the experience itself. The figure below showcases an example on how the virtual boundary looks like in the VR headset and how it changes red when the player crosses the boundary. Reason for this choice comes out of convenience provided by the hardware used in the study, but also to ensure that the boundary would be the same no matter the testing conditions allowing this study to be replicated more easily.

Figure 3.3 - Visual representation of the Oculus virtual boundary being crossed. Source: Meta, 2025, https://www.meta.com/nl/en/quest/safety-center/

3.2.2.1 Markers Development

This research uses the "Echoes of Abandonment" as a case study to research the influence of markers on the narration immersion. This research is concerned particularly with the diegetic and non-diegetic markers design. During the development phase the researcher has developed and tested 15 different ways to mark interactable object in the virtual space using Unity (all tests can be found in Appendix D). Those designs were based on the current standards of the industry, feedback from play testers, as well as theoretical background. These included variations such as outline glow effects, particle systems, colour tinting, floating icons of various styles, and proximity-triggered animations.

The selection criteria were grounded in Norman's (1988) concept of affordance and signifiers (mentioned in the chapter 2.2.1). The chosen markers needed to represent clear, distinguishable examples of their respective categories while maintaining functional equivalence in communicating interactability. Diegetic signifiers (e.g. bright toys suggesting "pick me up," brightened objects suggesting importance) communicate affordances through environmental integration - they function like real-world attention cues. Non-diegetic signifiers impose an external sign system that clearly communicates "this is interactable" but adds a layer of symbolic mediation between player and world. Additionally, the markers needed to meet the usability criteria, not to break the immersion by exposing the player to an unrecognisable design.

Several designs were rejected for failing one or more criteria. Outline glow effects, while commonly used in commercial games, created an ambiguous middle ground - participants in early testing were uncertain whether the glow was "part of the magic" or "the game showing me what to do." This semiotic ambiguity would have compromised the experimental design's requirement for clearly distinguishable conditions. Particle systems, though visually striking, drew excessive attention and risked becoming a confounding variable by differentially affecting immersion through visual spectacle rather than semiotic properties. Floating arrow labels were rejected for being too functionally different from the brightness condition - they provided explicit information rather than simply drawing attention.

The brightening shader was selected as the diegetic marker because it maintains the naturalistic-sensory coding orientation established by the clay-sculpted aesthetic of "Echoes of Abandonment." Light and luminescence are phenomena that exist within the physical world and can be plausibly interpreted as part of the child's imaginative perception - toys that "call out" to be played with through a gentle radiance. This aligns with what Fagerholt and Lorentzon (2009) describe as a diegetic interface element: one that exists both within the narrative fiction and within the game's spatial geometry. Brightened toys are visually standing out compared to the rest of the environment making them recognisable and signalising interaction possibility.

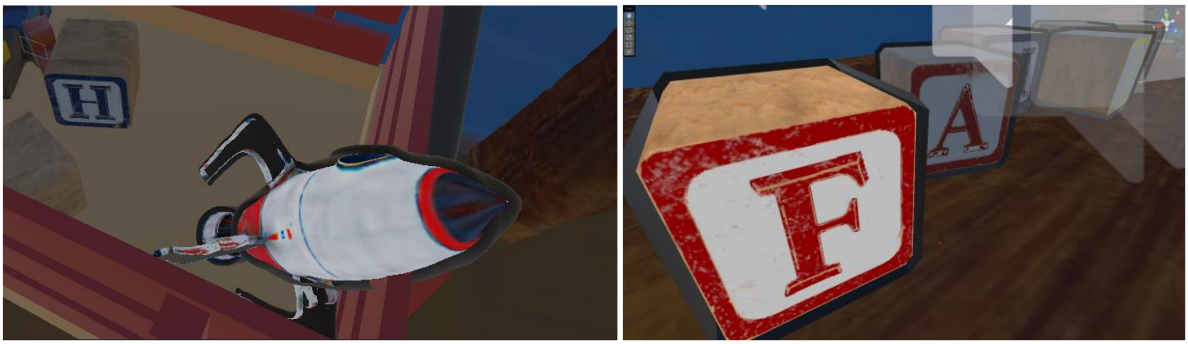

Figure 3.4 - Brightening as a diegetic marker - "Echoes of Abandonment" by Paulina Cywoniuk, 2024

The pulsating exclamation mark was selected for the non-diegetic condition precisely because it operates from a technological-abstract coding orientation, carrying strong conventional associations with game UI systems (Fagerholt & Lorentzon, 2009). This graphical element is aesthetically cohesive with the virtual environment, as an exclamation mark in a yellow circle it is styled to look like kid's crayon drawing. This design decision was informed by the theoretical distinction between an element's semiotic classification and its stylistic integration. The marker still operates as what Barthes (1977) would describe as a symbolic sign requiring learned interpretation, as players must understand the convention that an exclamation mark signals interactability.

Figure 3.5 - Graphical exclamation mark marker that represents the non-diegetic marker - "Echoes of Abandonment" by Paulina Cywoniuk, 2024

During iterative testing with 5 participants, the brightness was consistently described as feeling "natural" and "part of the world,", or simply "liked the most"; while the exclamation mark was readily identified as a "game element" or "hint system." This perceptual distinction is essential for the experimental design, as the research question requires markers that are clearly differentiated in their semiotic properties while remaining equivalent in their functional purpose of communicating affordances.

In relation to semiotics, diegetic markers function as *motivated signs* - they communicate through natural perceptual properties (brightness drawing the eye), while the non-diegetic markers function as *conventional signs* - they require learned associations and remind players they're in a game. Lastly, the control group (no markers) has no signifiers at all, but this control group might reveal some interesting data nonetheless, as well as creating the baseline of how the narration immersion is perceived in the "Echoes of Abandonment".

3.3 PARTICIPANTS

Participants of this study have been selected using a mix of convenience sampling and purposive sampling. Convenience sampling was used due to practical constraints of accessibility and time limitations of this master's thesis project, allowing for efficient recruitment of participants who were available. Purposive sampling was applied to ensure participants met specific criteria relevant to the research objectives, particularly regarding age range, location (research is based in the Netherlands) and physical mobility that would allow them to play a VR game while standing. This combined sampling approach is common in VR research conducted within academic settings, where practical accessibility must be balanced with the need to recruit participants who represent the target user population for the designed experience (Chiang, 2021; Köhle et al., 2021).

The sample size was determined by the practical constraints of the four-week data collection period available within the master's thesis timeline, combined with the logistical demands of individual VR testing sessions, as each requiring dedicated physical space, one-on-one supervision, and specialised hardware. Within these constraints, 35 participants were recruited and randomly assigned across three conditions (n = 11–12 per group). While this sample size limits the study's ability to detect small effects, it is consistent with comparable VR

research conducted under similar academic constraints. Köhle et al. (2021) investigated diegetic versus non-diegetic UI in VR with 37 participants across three conditions, and Saling et al. (2021) examined VR narrative engagement with 13 participants. The sample was sufficient to detect medium-to-large effects, which are the effect sizes most likely to carry practical significance for VR design decisions.

Participants were required to meet the following criteria: adults between 19 and 39; currently residing in the Netherlands; sufficient physical mobility to engage with VR hand controllers while standing; ability to speak, understand and read English. The age range of 19-39 was selected because this demographic represents a significant portion of VR game users (Kemp, 2025) and is likely to have baseline familiarity with digital gaming interfaces (Petri et al., 2020); they most likely balance technological familiarity with the emotional maturity to engage with the narrative themes of the experience (Arnett, 2000); and previous research on VR gaming and immersion has predominantly sampled from this age range (Köhle et al., 2021).

Due to the nature of testing in the VR, additional safety criteria have been implemented to ensure the comfort of the participants. The criteria exclude participants with history of photosensitive epilepsy or seizure disorders; severe motion sickness; or visual impairments that cannot be corrected with glasses or contact lenses compatible with the VR headset.
Since participants that have the VR headset on don't see the real-world surroundings and this standing experience might cause confusion among some participants, a physical and virtual border will be implemented in the testing location to ensure that the participants won't hurt themselves during the experience.

Participants were recruited through word of mouth and announcement on Teams at Breda University of Applied Sciences, Royal Academy of Art, the Hague, as well as personal and professional connections to the researcher. Recruitment materials described the study as involving a VR gaming experience followed by a questionnaire, with an estimated time commitment of approximately 20 minutes (10min game time and 10 min questionnaire). No compensation was given to the participants to support the lack of bias in their answers. Potential participants who expressed interest were screened against the inclusion and exclusion criteria before being scheduled for a session. All participants provided informed consent prior to participation.

3.4 RESEARCH PROCEDURE

The research procedure starts with participants individually entering a room with minimal amount of noise. The data has been collected from various locations, but it was ensured the conditions were the same to minimise unwanted variables. Rooms were around 5-10 m^2, and only the researcher and participant present. They would be asked to fill out the consent form, that can be found in the Appendix C. Upon completion, the researcher would install the VR headset on participant's head and turn on the game. Each participant is at random put in one of the 3 conditions of this study. The random allocation has been used by a quasi-systematic allocation with researcher discretion.

As the game starts, the first 10 seconds of the experience are the manual on how to move, interact and pause "Echoes of Abandonment". In case of any questions the researcher will be present next to the participant to help during this phase. After the participant shows familiarity of the interaction with the virtual space, the researcher will back away and let them experience the game without interruptions. The full tutorial is present in the following figure:

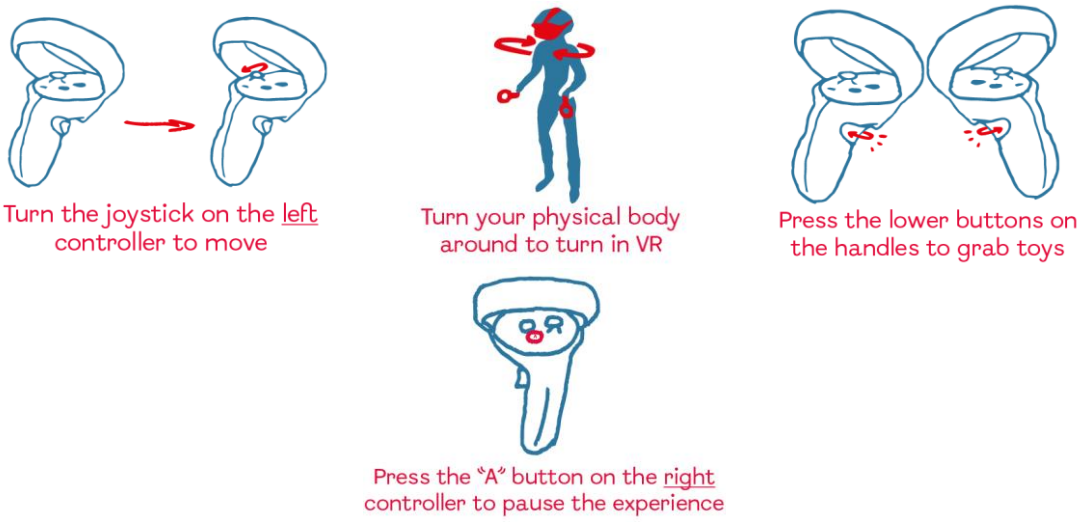

Figure 3.6 - Manual of "Echoes of Abandonment". "Echoes of Abandonment" by Paulina Cywoniuk, 2024.

Playing "Echoes of Abandonment" took about 10 minutes, but participants are free to finish the experience earlier if they wish. Upon completion of the game, the researcher will remove the headset from the participant and give them some time to sit down and rest their eyes after the standing VR experience. Afterwards, participant will be asked to complete the questionnaire presented to them by the researcher on the "Qualtrics" platform. This website has been chosen to collect data because of its history of usage in the quantitative research (Miller et al., 2020), as well as protection of the participants' data (Qualtrics (2025) - Security Statement. https://www.qualtrics.com/security-statement/ and Qualtrics (2025) - GDPR Compliance. https://www.qualtrics.com/gdpr/). The answers to the questionnaire will be stored in the cloud database of the chosen website to be retrieved again for the data analysis chapter.

The questionnaire was administered immediately after the VR experience on the same ASUS laptop used to run the Unity game build. After each participant completed the gameplay, the researcher closed the Unity application and opened the Qualtrics survey in a web browser, presenting the laptop to the participant to complete the questionnaire independently. This single-device setup ensured procedural consistency across all 35 sessions, though it meant participants transitioned directly from a standing VR experience to a seated laptop interaction, which was mitigated by the rest period described in the research procedure.

Participants are free to stay and have a conversation about the VR experience with the researcher in case they have some valuable input into the project, otherwise that completes the procedure for the participant. Additional time has been dedicated after each participant' procedure to ensure the lack of fatigue or any discomfort that the VR experience might have caused. The whole process is visualised in the flowchart below.

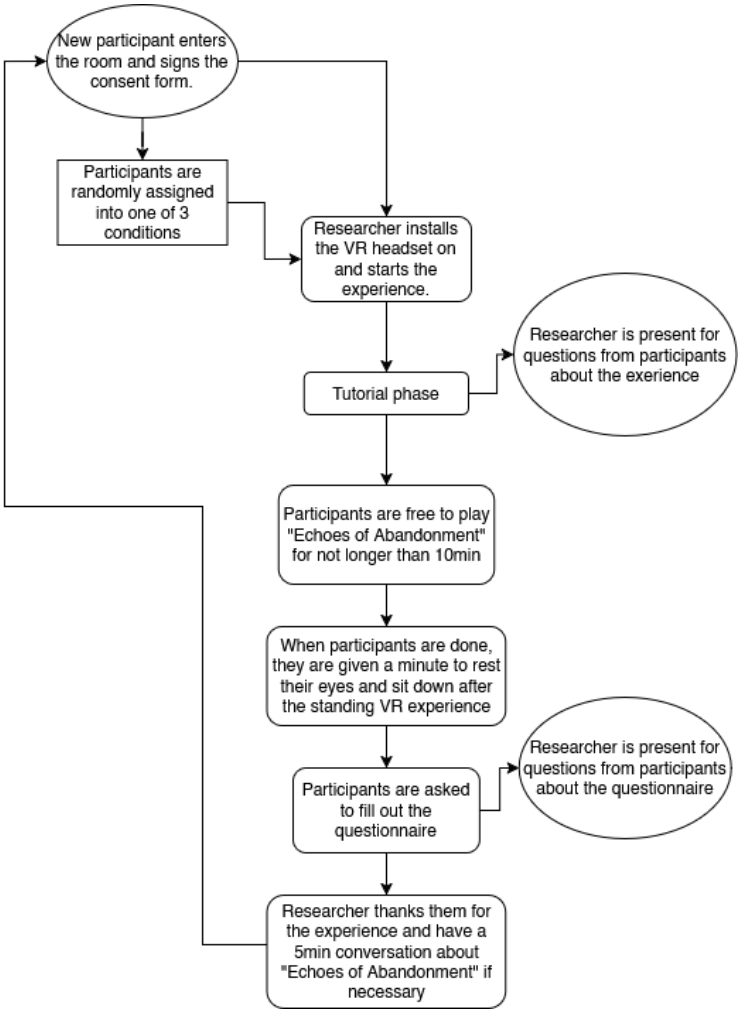

Figure 3.7 - Flowchart of the research procedure, Paulina Cywoniuk, 2025

3.5 TESTING METHODOLOGY

The methodology that was described in the previous chapter has been tested not only to ensure that the collected data and is analysis will be useful to prove the hypothesis, but also that the data collected that way will be the most optimal way to conduct this research. Pilot testing occurred in two phases: marker perception testing and full procedure testing. The participants of the pilot test were recruited by the researcher in the Breda University of Applied Sciences, and they met the criteria to take part in the study.

Participants in the first phase of the pilot test played "Echoes of Abandonment" for approximately ten minutes and provided feedback on the marker designs. This phase was essential for validating that the experimental conditions represented genuinely distinct semiotic categories. The diegetic marker (brightness shader) was consistently described as feeling "natural" and "part of the world," while the non-diegetic marker (exclamation mark icon) was identified as a "game element" or "hint system" making it a clear marker. This perceptual distinction confirmed that the markers successfully represented their respective categories - motivated signs integrated into the game world versus conventional signs requiring learned interpretation - validating their use in the main study. Crucially, no participant confused the categories - none described the exclamation mark as natural, and none identified the brightness

as a game interface element. This clear perceptual distinction validated that the markers successfully represented their intended semiotic categories: motivated signs integrated into the game world versus conventional signs requiring learned interpretation.

In the second phase, one pilot participant completed the full questionnaire, which revealed opportunities for refinement. Based on this feedback and subsequent discussions with thesis supervisors, two adjustments were made: a question about participants' previous VR experience was added to control for familiarity with the technology, and several questions were reformulated to improve clarity and comprehensibility. For instance the question 5 that now states "I found it easy to follow the story while playing the game", changed from "I was mentally involved in the narrative while playing the game", since "mentally involved" deemed to be too broad of a statement and it needed better vocabulary that wouldn't be so broad to interpret. Additionally, the number of questions has been reduced from 12 to 9 to ensure participants comfort and to minimize the fatigue connected to physically demanding VR game experience. while answering a questionnaire that now consists of 32 questions.

No technical issues with the game or procedural problems were identified during pilot testing - the Meta Quest 3 performed consistently, the marker conditions displayed correctly, and the audio narrative played as intended. The results of the pilot testing have been evaluated with research experts and in case any risks were identified the procedure has been adjusted accordingly.

While additional pilot participants would have strengthened procedural validation, the pilot testing conducted was sufficient to confirm the central requirement of this study: that the experimental conditions are perceptually distinct in ways that align with theoretical predictions. The consistency of participant descriptions across all five marker-testing participants suggests this finding is strong.

Taken together, the pilot testing validated the experimental approach on three levels: first, the marker conditions were perceptually distinguishable in the ways predicted by the theoretical framework - diegetic markers were consistently interpreted as part of the game world and non-diegetic markers as interface elements, confirming that the independent variable would function as intended. Second, the questionnaire items were interpretable to participants without requiring clarification, and the refinements made following pilot feedback (reformulating Question 5, reducing custom questions from 10 to 7) improved both clarity and completion comfort. Third, the procedural timing of approximately 10 minutes of gameplay followed by 15–20 minutes for the questionnaire was confirmed as manageable without inducing visible fatigue, supporting the decision to administer the questionnaire immediately after the VR experience.

3.6 DATA ANALYSIS

The answers gathered by the participants filling out the questionnaire will be analysed using the analysis of variance (ANOVA) statistical test to prove the influence of the independent variable (markers) and the dependent variable (narration immersion). One-way ANOVA is the appropriate statistical test when comparing means across three or more independent groups with a continuous dependent variable (Field, 2018). This matches the present study's design: three independent groups (diegetic markers, non-diegetic markers, no markers) and continuous dependent variables (IEQ-SF and NES scores).

ANOVA has been favoured over alternative statistical analysis framework such as the Kruskal-Wallis since it matched the design of this study strongly. ANOVA assumes the data collected in this study meets certain conditions: the dependent variable is continuous, the data is approximately normally distributed within each group, and variances across groups are roughly equal – as is expected in this study design. Kruskal-Wallis, on the other hand, makes no assumptions about distribution. It ranks all data points and compares median ranks rather than means. One-way ANOVA was selected over the non-parametric alternative, Kruskal-Wallis test, for several reasons. First, while individual Likert scale items are ordinal, the composite scores derived from summing multiple scale items (11 items for IEQ-SF, 12 items for NES) approximate continuous interval-level data, satisfying ANOVA's measurement requirements (Norman, 2010; Sullivan & Artino, 2013). Secondly, parametric tests offer greater statistical power than their non-parametric counterparts when assumptions are reasonably satisfied, increasing the likelihood of detecting true effects (Pallant, 2020).

After the data has been collected the researcher of this thesis will run a validity test to see if the questionnaire is valid. This will be especially important for the questions developed by the researcher for the purpose of this study, as they have not been academically validated yet, but also for the IEQ and NES since those scales have not been tested together yet. To do so the reliability analysis Cronback's alpha will be used. Cronbach's alpha measures the extent to which items within a scale measure the same underlying construct, with values above 0.70 generally considered acceptable (Tavakol & Dennick, 2011).

ANOVA is an appropriate statistical test to use in his study because it allows comparison of means across three independent groups: diegetic markers (Group A), non-diegetic markers (Group B), and no markers (Group C). Two separate ANOVA tests will be performed:

1. ANOVA comparing the three marker conditions on total IEQ-SF scores (measuring game immersion)
2. ANOVA comparing the three marker conditions on total NES scores (measuring narrative engagement)

The significance level will be set at $p < 0.05$, following standard conventions in social sciences research (Field, 2018). If the ANOVA reveals a statistically significant difference between groups, post-hoc tests using Tukey's (1953) Honestly Significant Difference (HSD) will be conducted to identify which specific group comparisons are significant. This step is essential for testing Hypothesis 1, which predicts that non-diegetic markers will disturb narrative immersion compared to other conditions.

3.6.1 Secondary analysis for the custom questions

The custom questions introduced by the researcher for the purpose of this study call for separate data analysis since those questions have not been academically validated unlike the used scales. This analysis will be done descriptively, reporting means and standard deviations for each experimental condition. Descriptive statistics provide an overview of the data that supports interpretation of inferential tests (Pallant, 2020). These questions serve to support interpretation of the primary scale results by providing insight into specific aspects of the experience such as perceived usability of markers (Question 4) and relationship between markers and story comprehension (Question 6).

Two questions from this section of the questionnaire are not on the standardized Liker scale. Question 1 will be reported descriptively to characterise the sample and may be examined as a potential covariate if distribution across conditions is uneven. Responses to Question 7 (open-ended question about marker design) will be analysed thematically to identify recurring patterns in participant perceptions. Qualitative data complements quantitative findings by providing rich contextual information that helps explain why certain patterns emerged (Lankoski & Bjork, 2015). These qualitative insights will help contextualize the quantitative findings and may reveal dimensions of marker perception not captured by the structured questions.

3.6.2 Validity of data collected

Ensuring the validity and reliability of data collection is essential for producing meaningful research findings. This section outlines the measures taken to ensure that the methodology of this study produces data that accurately measures the constructs under investigation.

In this study, narrative immersion is operationalized as the combination of two constructs: immersion experience (measured by IEQ-SF) and narrative engagement (measured by NES). This operationalization is grounded in the theoretical framework established in Chapter 2, which argues that narrative immersion in VR games requires both absorption in the game environment and engagement with story content - neither construct alone fully captures the phenomenon.

Both scales selected for this study have been academically validated in prior research. The Immersion Experience Questionnaire Short Form (IEQ-SF) was validated by Cutting et al. (2025) as a reliable measure of game immersion. The Narrative Engagement Scale (NES) was developed and validated by Busselle and Bilandzic (2009) for measuring experiential engagement with narratives. By combining these established instruments, this study builds on validated measurement approaches rather than relying solely on untested measures.

The custom questions developed for this study serve as control variables to strengthen interpretation of results. For instance, Question 2 verifies that baseline usability did not differ between conditions, helping to rule out the alternative explanation that differences in narrative immersion resulted from usability problems rather than marker design.

Prior to hypothesis testing, Cronbach's alpha will be calculated for both the IEQ-SF and NES scales to verify internal consistency within this study's sample. Values above 0.70 are generally considered acceptable (Tavakol & Dennick, 2011). This step is particularly important as these two scales have not previously been used together in VR narrative research.

Additionally, pilot testing with five participants was conducted prior to the main study to identify potential issues with the questionnaire and research procedure. This iterative testing helped refine question formulations and confirmed that the methodology functions as intended.

3.7 LIMITATIONS

Despite the best efforts of the researcher of this thesis, this study has some limitations that need to be acknowledged.

The custom questions developed for this study have not been academically validated, which may affect the reliability of data collected through these items. To mitigate this concern,

the custom questions were designed to supplement rather than replace the validated scales (IEQ-SF and NES), and their formulation was refined based on pilot testing and supervisor feedback. However, if results show differences between conditions, the current questions may not provide sufficient data to fully explain the underlying mechanisms, requiring some interpretation to rely on theoretical reasoning rather than direct measurement.

Sampling chosen to conduct this study has some limitations. This study used a combination of convenience and purposive sampling, which introduces potential biases that may affect the generalizability of findings. Convenience sampling, while practical given the time constraints of this master's thesis, means participants were selected based on accessibility to the researcher rather than representativeness of the broader population. As recruitment occurred primarily through academic networks at Breda University of Applied Sciences and the Royal Academy of Art in The Hague, the sample likely over-represents individuals with higher education levels and greater familiarity with research participation. Future research should employ probability sampling methods and broader recruitment strategies to enhance the external validity of findings in this area.

Additionally, measuring narrative engagement in interactive media presents inherent challenges. According to Qin, Rau, and Salvendy (2009), narratives in games differ substantially from traditional linear narratives in other media, and therefore "most measurements for traditional narrative are not adequate for analysing the computer game narrative." While the Narrative Engagement Scale (Busselle & Bilandzic, 2009) was developed for screen-based narratives rather than interactive VR, it remains one of the most established instruments for measuring narrative engagement. This limitation is partially addressed by combining it with the IEQ-SF, which was specifically designed for game contexts.

Despite theoretical grounding and pilot testing validation, the marker designs may not produce effects significant enough to detect with the sample size of this study. During pilot testing, participants clearly distinguished between the diegetic marker (described as "natural" and "part of the world") and the non-diegetic marker (identified as a "game element"), suggesting the designs successfully represent their intended semiotic categories. Furthermore, the marker selection was grounded in semiotics (Saussure, 1917) and the diegetic framework (Fagerholt & Lorentzon, 2009), providing theoretical support for why these specific designs should differentially affect narrative immersion. Nevertheless, the possibility remains that the effect size may be smaller than anticipated.

As much as the case study of this research has been selected for variety of positive reasons it might prove limiting at the same time. Firstly, The narrative of "Echoes of Abandonment" is presented entirely through audio - parents' voices arguing behind a door - without subtitles or visual text. This design choice means that participants must be able to hear clearly and the testing environment must have appropriate acoustics. Consequently, participants with hearing impairments cannot fully access the narrative experience. Due to time constraints, adding subtitles was not feasible for this study; future research should consider implementing accessible alternatives to ensure broader participant inclusion.

The target group of this study (adults aged 19-39, residing in the Netherlands, able-bodied, English-speaking) does not represent the general population. This age range was selected because it represents a significant portion of VR game users and aligns with previous VR immersion research (Köhle et al., 2021). However, findings may not generalize to younger

players, older adults, or populations in other geographic regions. Ideally, this research would sample from a broader demographic; however, the timeframe of this master's thesis project and accessibility constraints limited recruitment to participants within the Netherlands and the researcher's academic networks.

Administering questionnaires immediately after a VR experience risks participant fatigue, which may affect response quality. VR experiences can be physically and cognitively demanding, particularly for participants unfamiliar with the technology. To minimize this risk, the shortened version of the IEQ (IEQ-SF) was specifically selected to reduce questionnaire length, and the total questionnaire was designed to take no longer than 15-20 minutes to complete. Additionally, participants are given time to rest after removing the headset before beginning the questionnaire.

In conclusion, the limitations of this study have been acknowledged, and a prevention plan has been created to pivot from them. In the Chapter 4 the full methodology will be used to collect the data and analyse them in a meaningful way to answer the Hypothesis 1 and 2.

3.8 ETHICAL CONCERNS

The research of this thesis is a subject to certain ethical considerations that need to be addressed.

Firstly, as mentioned before, this study is conducted in the VR headset designed to be used in the standing position with audio-based story that requires no hearing problems. This form of experience can be difficult for people with disabilities, motor limitations or other conditions. The researcher considered several alternatives to broaden accessibility: implementing subtitles or visual text cues to supplement the audio narrative, adding post-hoc captioning to the parents' dialogue, and redesigning the narrative delivery to include a visual storytelling component. However, each of these alternatives would have fundamentally altered the sensory conditions of the experiment. Since this study investigates the influence of visual markers on narrative immersion, introducing additional visual narrative elements, such as subtitles overlaid on the game environment, would have introduced a confounding variable by adding visual information that could interact with the marker conditions in unpredictable ways. The decision to maintain an audio-only narrative and consequently exclude hearing-impaired participants was therefore made to preserve the internal validity of the experimental design, rather than being a simple oversight.

Ethical considerations of this study not only concern themselves with the physical health of the participants, but also their mental health. The nature of the story on which the research was conducted might be triggering for some people who have uncomfortable memories of their parents' argument. To prevent this each participant has been handed a consent form that described the possibility of having uncomfortable experience. Additionally, to informing participants beforehand of the subject of the experiment, they were given the opportunity to withdrawal from the study at any point of time without consequences.

The researcher of this study was present during the gameplay to ensure the technical support during the study. Unfortunately, this might create social desirability effects or pressure to continue when uncomfortable towards the participants. To mitigate this, the procedure includes the researcher to quietly observe the play without interruption, with exception of the

manual phase and the questionnaire phase. It is expected that the participants' immersion would cause them to focus on the game rather than on the presence of the researcher in the room. Additionally, in the consent for the participants will be informed of the possibility to withdrawal from the study at any point, even if they feel any discomfort.

Beyond the immediate ethical considerations of the research procedure, it is important to acknowledge the ethical implications of the participant selection for the data this study produces. Recruitment occurred primarily through academic networks at Breda University of Applied Sciences and the Royal Academy of Art in The Hague, as well as a game developers' event. This means the sample disproportionately represents individuals with higher education and, in many cases, training in design, game development, or related creative fields. Participants with this background may perceive and interpret visual markers, such as the diegetic brightness shader or the non-diegetic exclamation mark, differently from the general population, as their professional training may have sensitised them to UI conventions or made them more forgiving of interface elements that a general audience might find disruptive. This raises an ethical question about whose experiences this research represents: the findings reflect the perspectives of a predominantly young, educated, design-literate population in the Netherlands, and may not capture how markers are experienced by players without this background. This consideration informed the inclusion of Question 8 (profession/specialty) in the questionnaire, which allows the researcher to examine whether professional background influenced responses, but it does not fully resolve the representational limitation.

Lastly, one of the concerns that this research might pose is the secure storage of participants'' data. Answers to the survey will be collected through the Qualtrics platform, which was selected for its secure data storage practices and compliance with research ethics standards. All responses will be collected anonymously - no personally identifiable information (such as names or email addresses) will be linked to questionnaire responses. Only the researcher and thesis supervisors will have access to the raw data. Participants retain the right to request deletion of their data at any point during or after the study, as they will be informed about it in the consent form. The data will be retained until the thesis is completed and graded, after which it will be securely deleted within 30 days of final thesis submission and grading. Any data referenced in the published thesis will be presented in aggregate form only, ensuring individual responses cannot be identified.

3.8.1 Researcher bias – influence of qualitative analysis

The researcher of this study occupies a dual role as both the creator of "Echoes of Abandonment" and the investigator examining player responses to it. Having designed, developed, and publicly exhibited the game as a bachelor's graduation project, the researcher brings both intimate knowledge of the experience and a personal investment in how it is received. This positionality must be acknowledged, as it introduces potential biases at multiple stages of the research process.

During data collection, the researcher was physically present during all testing sessions and participants were aware that they were playing a game made by the person conducting the study. This awareness may have introduced social desirability bias, as participants could have

been more positive in their feedback, less willing to express criticism of the game, or more inclined to continue with the experience even if uncomfortable. The informal post-session conversations, in which participants shared their thoughts verbally with the game's creator, are particularly susceptible to this effect. Additionally, because verbal comments were noted from memory shortly after each session rather than audio-recorded, the qualitative data reflects the researcher's recollection and interpretation rather than verbatim participant accounts.

During data interpretation, the researcher's familiarity with the game may have influenced the qualitative analysis - specifically, the thematic coding of open-ended responses, the selection of which observations to foreground, and the framing of participant behaviour. A researcher who built the experience may be more attuned to patterns that align with their design intentions and less attuned to those that contradict them.

Several measures were taken to mitigate these risks. The primary data source - the questionnaire - was administered through the Qualtrics platform rather than verbally, reducing interviewer effects on the quantitative responses. The validated scales (NES and IEQ-SF) and the statistical analysis pipeline (conducted in SPSS following standard assumption testing procedures) are reproducible and not subject to researcher interpretation. Random assignment to conditions was ensured through a random number generator rather than researcher judgement. For the qualitative analysis, Braun and Clarke's (2006) thematic analysis framework was followed to provide systematic structure to the coding process. Throughout the project, thesis supervisors reviewed analytical decisions and provided external accountability on the interpretation of findings. These safeguards do not eliminate researcher bias, but they constrain its influence to the qualitative components of the study, where it is explicitly acknowledged in Chapter 4.

4 DATA AND RESULTS

The purpose of this study is to investigate the research question "*What is the influence of different markers design in VR games on the player's ability to become immersed in narrative content among young adults living in the Netherlands?"* by testing the hypothesis 1 or/and 2, posed in chapter 3.4.3. The methodology of this research has been executed as described in the previous chapter, following the analysis pipeline of Field (2018) for ANOVA assumptions, which provided the researcher with results that will be analysed in the following chapters.

4.1 DATA

A total of 35 participants were recruited and randomly assigned to one of three experiment conditions: diegetic markers (n = 12), non-diegetic markers (n = 11), and control group with no markers (n = 12). Participants were assigned to conditions using a quasi-systematic rotation method. The researcher aimed to distribute participants evenly across the three conditions by following an approximate rotating sequence (control, diegetic, non-diegetic), with occasional deviations from this order based on practical scheduling considerations. The final sample size was determined based on practical constraints of the thesis timeline. All participants were aged between 19 and 39 as specified in the inclusion criteria. No participants withdrew from the study nor stopped the VR experience prematurely. All participants signed the consent form before starting the study and were briefed with the same prompt described in the Appendix E.

Data was collected across four types of locations: university classrooms at Breda University of Applied Sciences and Royal Academy of Art in the Hague, the researcher's home, and a private room rented during a game developers' event in The Hague. The researcher acknowledges that collecting data in those locations/events might have attracted specific demographic (mostly game related students/professionals) - the limitations of this convenience sampling will be mentioned in the chapter 4.5.
While the locations varied, the researcher ensured consistent conditions across all settings: each space was quiet and provided sufficient room for the standing VR experience with the Meta Quest 3's virtual boundary system. The use of the same VR hardware, game build, and questionnaire platform (Qualtrics) across all sessions ensured procedural consistency despite the varying physical environments. The potential influence of environmental differences on participant responses is acknowledged as a limitation of this study.

One notable disruption occurred during the session of participant 5 (diegetic condition), when several individuals entered the room during gameplay, creating significant noise disturbance. This participant's data was retained in the analysis, as excluding a single data point from an already small sample would reduce statistical power further. However, this disruption could have affected this participant's immersion and engagement responses, that's why report for this participant is stated here. After running the analysis, it has been noted that participant 5's scores were within the normal range for the diegetic group, therefore the data from this participant will remain.

No technical issues with the Meta Quest 3 hardware, game performance, or questionnaire administration were encountered during data collection. All 35 participants completed both the

VR experience and the full questionnaire. Only participant 24 (assigned to non-diegetic condition) missed one question in the survey – namely IEQ-SF_4 "I enjoyed playing the game ". The analysis proceeded despite this missed answer.

The qualitative data presented in chapter 4.3, should be interpreted with awareness of the researcher's dual role as both game developer and interviewer. Participants were aware that the researcher had created 'Echoes of Abandonment,' which may have influenced their willingness to offer critical feedback. The verbal comments were gathered through informal post-session conversations rather than structured interviews, and no audio recording was made - the researcher noted key comments from memory shortly after each session. These methodological choices mean the qualitative findings represent the researcher's interpreted reconstruction of participant feedback rather than verbatim accounts, and should be weighted accordingly.

4.1.1 Level of experience with VR games

One of the open answer questions asked to the participants of this study in the questionnaire was their statement of experience with the VR games. The total demographic consisted of 48.6% participants having no prior experience with VR games, 34.3% reporting low level of VR experience, 8.6% reporting mid-level of experience and 8.6% with high or pro level of VR experience. Since the questions was open ended the researcher has done this classification independently based on the self-reported answers. As mentioned before, distribution to the three conditions was random. The following graph represents how the final distribution looked like in relation to the level of experience with VR and assigned condition.

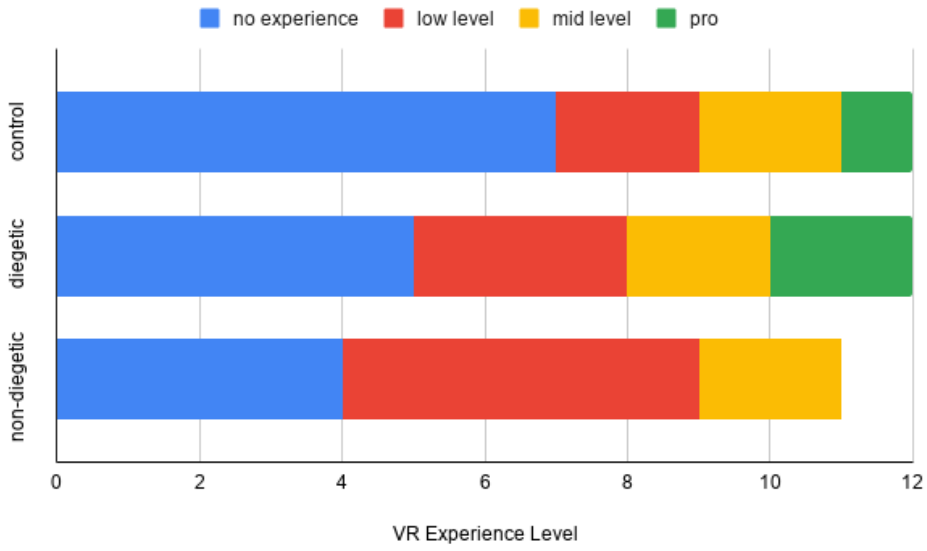

Figure 4.1 - Distribution of players with self-reported VR experience level in each condition.

4.1.2 Profession of participants

One of the open answer questions that was asked to the participants in the questionnaire regarded their profession/specialization to ensure that not only the demographic varied, making this study's results applicable to larger population, but also to investigate the influence of game-proficient participants on their perception of immersion narrative, compared to not-game-experienced participants.

Participants came from diverse professional and academic backgrounds. Based on self-reported professions and specialisations, the sample included participants from game-related fields such as game design, game art, game programming, sound design, and UI/UX design (12 participants, 34.3%); creative fields such as graphic design, photography, and creative business (8 participants, 22.9%); technology-related fields such as software engineering and development (4 participants, 11.4%); and other non-technical fields such as literature, urban planning, built environment, tourism, and psychology (11 participants, 31.4%). The name of those categories involved researcher judgement, as several responses spanned multiple disciplines (for example, a participant studying built environment who also holds a psychology degree). This division into four categories is therefore subjective, but nonetheless interesting to investigate in case there are any statistical relationships between participants' specialization and their ability to recognise interactable objects in virtual spaces. Pie chart below represents the stated values in a visual format.

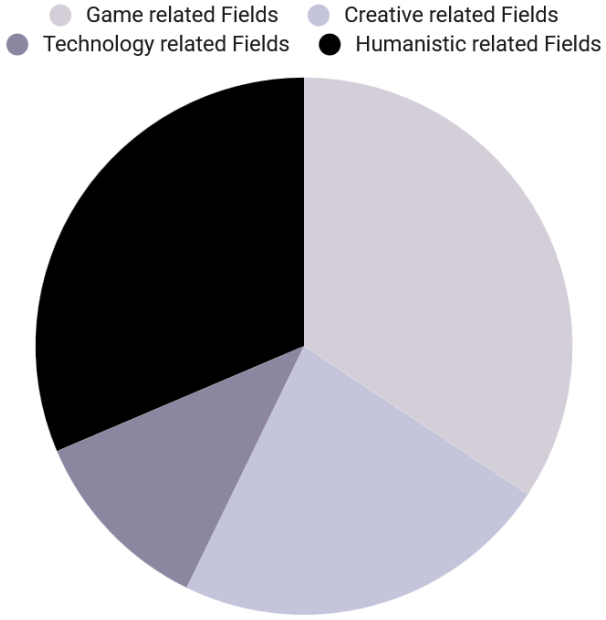

Figure 4.2 - Pie chart representing participants' professions

4.2 QUANTITATIVE ANALYSIS

The main source of data for this project came from the answers to the questionnaire that the participants were asked to fill out after having completed the 7-10min experience in VR. The answer from the survey will be analysed using the statistics leading towards ANOVA analysis. All analysis was conducted in SPSS 31 (newest version in the time or writing this thesis).

4.2.1 Scale Reliability

Before conducting the main analyses, Cronbach's alpha was calculated for both scales and their subscales to assess internal consistency. This value indicates reliability that defined as "how accurately a test measures the thing which it does measure" (Kelley, 1927). This step was particularly important given that the IEQ-SF and NES have not previously been used together in VR narrative research, and the IEQ-SF was originally validated for screen-based gaming rather than immersive VR experiences. All items requiring reverse scoring were recoded prior to analysis (IEQ-SF items 5 and 10; NES items 1–6).

4.2.1.1 Narrative Engagement Scale (NES) reliability

The NES demonstrated strong overall reliability (α = .869), well above the .70 threshold generally considered acceptable (Tavakol & Dennick, 2011). Three of its four subscales also showed good to excellent reliability: Attentional Focus (α = .823), Narrative Presence (α = .839), and Emotional Engagement (α = .805). The Narrative Understanding subscale fell below the acceptable threshold (α = .624). This lower reliability may reflect the VR context - in a virtual environment, "making sense of what was going on" could relate to understanding the VR interface itself rather than the narrative, introducing measurement noise into this subscale. Possibilities for this score will be presented in the Discussion chapter. Based on these results, the NES total score and the three reliable subscales (Attentional Focus, Narrative Presence, Emotional Engagement) were used as the primary dependent variables in this study.

4.2.1.2 Immersion Experience Questionnaire – Shortened Version (IEQ-SF)

The IEQ-SF showed poor overall reliability (α = .487), substantially below acceptable thresholds. At the subscale level, Involvement (α = .659) and Challenge (α = .650) approached but did not reach the .70 threshold, while Real-World Dissociation performed poorly (α = .439). Therefore, the reliability of the IEQ-SF and its subscales was lower than acceptable threshold for this research's sample.

This poor reliability constitutes a methodological finding in itself: the IEQ-SF's factor structure, developed for screen-based gaming (Cutting et al., 2025), may not transfer to narrative VR experiences where the headset itself provides real-world dissociation by default, what questions the statement of the Cutting et al. 2025 paper where they say "(...) IEQ-SF is a good substitute for the full IEQ in both high and low immersion situ-ations.". This theme will be analysed more in the Discussion chapter (5.2).

Despite the reliability limitations, the IEQ-SF Involvement and Challenge subscales were retained for exploratory analysis with the caution that results should be interpreted critically. This decision was made to provide a more complete picture of the data rather than discarding collected responses entirely.

The table below summarises the reliability results across both scales.

Scale / Subscale	**Cronbach's α**	N of items	Status
NES Total	**.869**	12	Excellent
Attention Focus	.823	3	Good
Narrative Presence	.839	3	Good
Emotional Engagement	.805	3	Good
Narrative understanding	.624	3	Below threshold
IEQ-SF Total	**.487**	11	Poor
Involvement	.659	4	Below threshold
Real-World Dissociation	.439	4	Poor
Challenge	.650	3	Below threshold

Table 4.2 - table of α values across scales and subscales used in this research, following number of items and status of said α value.

4.2.2 Descriptive overview

The mean responses per item across the three conditions are presented in Figure 4.3. This visualisation reveals that participants in all three conditions responded remarkably similarly across the majority of items, with most narrative engagement items scoring between 4 and 6 on the 7-point scale, indicating moderate to high engagement regardless of marker condition.

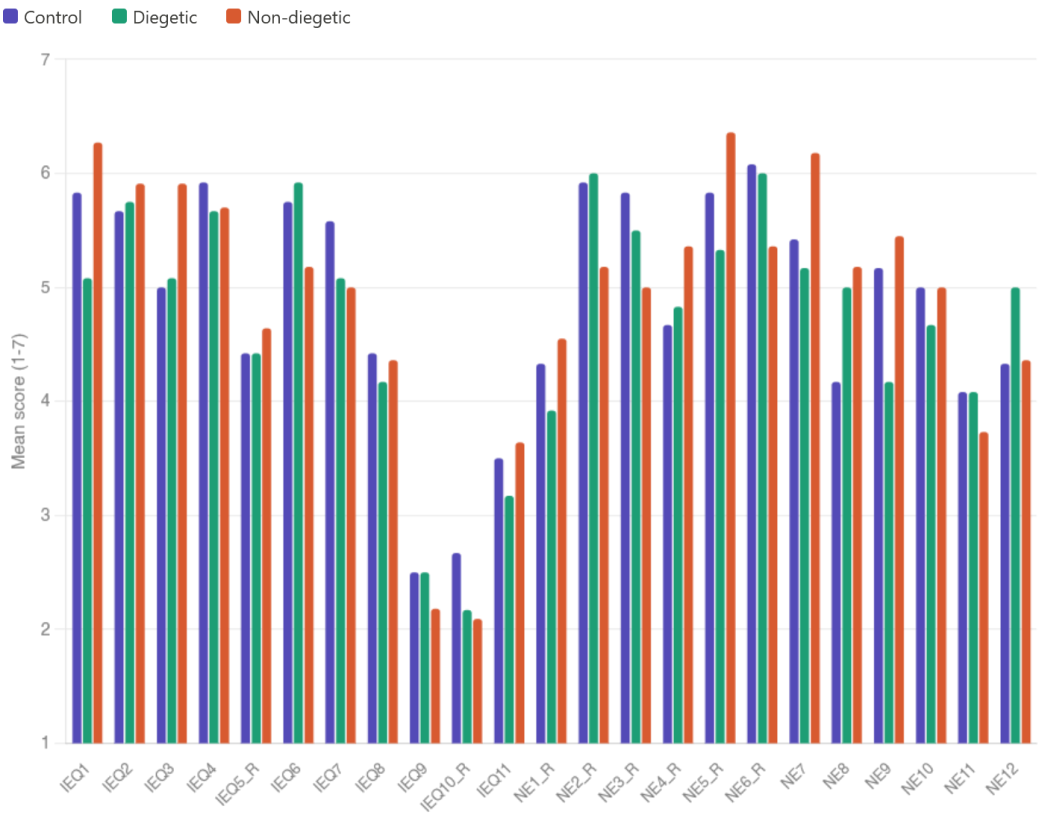

Figure 4.3 - Graph with mean score of all elements of NES and IEQ-SF across three conditions.

The most notable pattern in the item-level data is the consistently low scores on the IEQ-SF Challenge items (IEQ9, IEQ10_R, IEQ11), where all groups scored between approximately 2 and 3.5. This confirms that participants did not find "Echoes of Abandonment" particularly challenging, which is consistent with its design as a narrative exploration experience rather than a skill-based game.

4.2.3 Assumption testing

Prior to the main inferential analyses, the assumptions of normality and homogeneity of variance were tested for each dependent variable. One-way ANOVA is a parametric test that assumes the data within each group is approximately normally distributed and that the variances across groups are equal (Field, 2018). When these assumptions are violated, the test may produce unreliable p-values, particularly with small sample sizes. Verifying these assumptions determines whether parametric (ANOVA) or non-parametric (Kruskal-Wallis) tests are appropriate for each scale/subscale.

4.2.3.1 Outliers

Boxplot inspection during assumption testing revealed several mild outliers across the dependent variables: NES Attentional Focus (participants 9 and 18 in the non-diegetic group),

NES Narrative Presence (participant 34 in the control group, participant 17 in the non-diegetic group), NES Emotional Engagement (participant 4 in the diegetic group), and IEQ Challenge (participant 20 in the non-diegetic group). All outliers were mild (within 1.5 to 3 times the interquartile range) rather than extreme. These data points were retained in the analysis for two reasons: first, with only 11–12 participants per group, removing data points would substantially reduce statistical power; second, no methodological grounds were identified to justify exclusion, as these responses represent natural variation in participant experience rather than measurement errors (Field, 2018). The potential influence of these outliers on the results is acknowledged in the Discussion chapter.

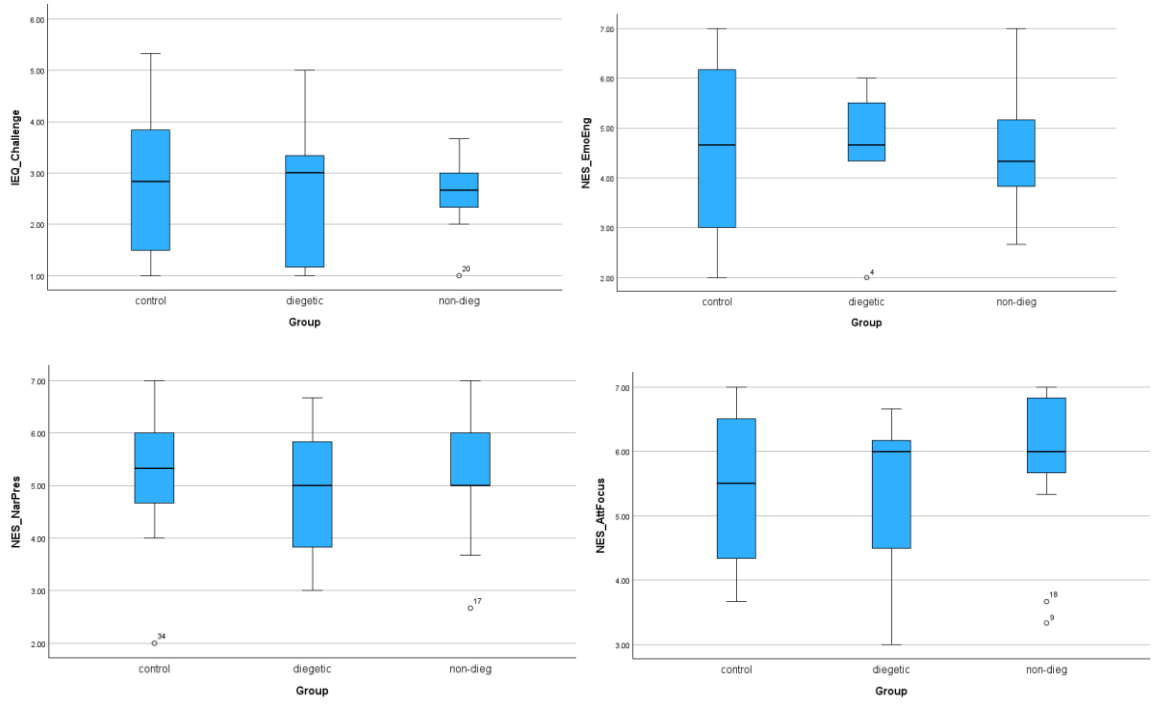

Figure 4.4 - Box plots of outliers per subscale and condition – from left up to right down IEQ_Challange, NES_EmoEng, NES_NarPres, NES_AttFocus

4.2.3.2 Normality

The Shapiro-Wilk test was used to assess normality for each dependent variable within each condition. Normality was satisfied ($p > .05$) for NES Total (control group $p = .995$, diegetic $p = .311$, non-diegetic = .801), NES Narrative Presence (control group $p = .312$, diegetic $p = .844$, non-diegetic = .435), IEQ Involvement (control group $p = .467$, diegetic $p = .525$, non-diegetic = .518), and IEQ Challenge (control group $p = .644$, diegetic $p = .206$, non-diegetic = .281) across all three groups. Normality was violated for NES Attentional Focus (control group $p = .405$, **diegetic p = .038**, **non-diegetic = .017**) and NES Emotional Engagement (control group $p = .335$, **diegetic p = .023**, non-diegetic = .780).

Consequently, one-way ANOVA was used for variables meeting the normality assumption, while the non-parametric Kruskal-Wallis test was used for Attentional Focus and Emotional Engagement. Data transformation (e.g., log transformation) was considered as an alternative to preserve parametric testing, but given the small sample size and the availability of an appropriate non-parametric alternative, the Kruskal-Wallis test was selected as the more conservative approach. While Norman (2010) argues that parametric tests are robust to violations of normality with summed Likert scales, this robustness is less certain with small samples of 11-12 per group and individual subscales comprising only three items.

4.2.3.3 Homogeneity Variance

Levene's test was conducted for all variables analysed with ANOVA to verify homogeneity of variance. This assumption requires that the variability in scores is approximately equal across all groups - if one group's responses are much more spread out than another's, ANOVA's comparison of group means becomes unreliable (Field, 2018). Levene's test evaluates whether the variances differ significantly; a non-significant result ($p > .05$) indicates the assumption is met. Homogeneity of variance was satisfied for NES Total ($p = .266$), NES Narrative Presence ($p = 1.000$), IEQ Involvement ($p = .454$), and IEQ Challenge ($p = .055$). All variables therefore met this assumption for ANOVA.

4.2.4 Main analysis

All the above tests have been leading to the inferential analysis that this research needs to ensure conclusions. ANOVA or Kruskal-Wallis test has been used to assess the effect of markers conditions on different scales and subscales that met the reliability requirements.

4.2.4.1 NES Total ANOVA

A one-way ANOVA was conducted to compare the effect of marker condition on overall narrative engagement as measured by the NES Total score. The analysis revealed no statistically significant difference between the three conditions, $F(2, 32) = .063$, $p = .939$, $\eta^2 = .004$. The group means were nearly identical: control ($M = 5.15$, $SD = 1.15$), diegetic ($M = 5.01$, $SD = 0.69$), and non-diegetic ($M = 5.11$, $SD = 1.06$). Tukey's post hoc comparisons confirmed no significant pairwise differences (all $p > .05$). The effect size ($\eta^2 = .004$) indicates that marker condition explained less than 1% of the variance in narrative engagement scores.

This absence of difference across conditions is consistent with the findings of Köhle et al. (2021), who concluded that different UI approaches were suited to different contexts rather than one being inherently superior, though their study measured user preference in VR shooters rather than narrative immersion. The near-identical means across all three groups suggest that marker type may not be a meaningful predictor of narrative engagement in this context - a possibility explored further in the Discussion chapter.

4.2.4.2 NES Subscales Analyses

Attentional Focus: A Kruskal-Wallis test revealed no significant difference between conditions, $H(2) = 1.298$, $p = .522$. Median ranks were: control = 5.50, diegetic = 6.00, non-diegetic = 6.00. Despite reporting median for the Kruskal-Wallis results is a standard, for consistency with the ANOVA results, the means and standard deviation are hereby listed. Mean = 5.44 for control group, 5.36 for diegetic and 5.81 for non-diegetic condition, and SD control = 1.18, diegetic = 1.24, non-diegetic = 1.27.

Narrative Presence: A one-way ANOVA revealed no significant difference between conditions, $F(2, 32) = 0.327$, $p = .723$, $\eta^2 = .020$. Group means were: control ($M = 5.19$, $SD = 1.33$), diegetic ($M = 4.86$, $SD = 1.13$), non-diegetic ($M = 5.24$, $SD = 1.27$).

Emotional Engagement: A Kruskal-Wallis test revealed no significant difference between conditions, $H(2) = 0.452$, $p = .798$. Medians were: control = 4.67, diegetic = 4.67, non-diegetic = 4.33. Once again, for consistency with the ANOVA tests means and SD will be reported as well. Mean for control group was 4.61, SD 1.77, diegetic mean = 4.75, SD = 1.06, and for the non-diegetic mean = 4.48, SD was 1.29.

4.2.4.3 Exploratory Analysis: IEQ-SF Subscales

Given the reliability limitations discussed in section 4.2.2, the following results should be interpreted with caution.

Involvement: A one-way ANOVA revealed no significant difference between conditions, $F(2, 31) = 0.275$, $p = .761$, $\eta^2 = .017$. Group means were: control ($M = 5.60$, $SD = 0.97$), diegetic ($M = 5.40$, $SD = 1.04$), non-diegetic ($M = 5.68$, $SD = 0.72$).

Challenge: A one-way ANOVA revealed no significant difference between conditions, $F(2, 32) = 0.192$, $p = .826$, $\eta^2 = .012$. Group means were: control ($M = 2.89$, $SD = 1.41$), diegetic ($M = 2.61$, $SD = 1.31$), non-diegetic ($M = 2.64$, $SD = 0.77$).

4.2.5 Summary of inferential results

No statistically significant differences were found between the three marker conditions on any measure of narrative engagement or game immersion. All p-values were substantially above the .05 significance threshold, and all effect sizes were negligible ($\eta^2 < .02$).

Regarding Hypothesis 1 ("The diegetic markers will support the narration immersion"), Tukey's post hoc comparison between the diegetic group and the control group on NES Total showed no significant difference ($p = .936$), with the diegetic group scoring slightly lower ($M = 5.01$) than the control group ($M = 5.15$). This pattern was consistent across all subscales, with no measure showing significantly higher scores for the diegetic condition. Hypothesis 1 is therefore not supported.

Regarding Hypothesis 2 ("The non-diegetic markers will noticeably disturb the narration immersion"), Tukey's post hoc comparison between the non-diegetic group and the control group on NES Total showed no significant difference ($p = .995$), with the non-diegetic group scoring nearly identically ($M = 5.11$) to the control group ($M = 5.15$). No measure showed significantly lower scores for the non-diegetic condition. Hypothesis 2 is therefore not supported.

These findings indicate that neither the diegetic brightness markers nor the non-diegetic exclamation mark markers significantly influenced participants' narrative immersion compared to the control condition with no markers.

4.2.6 Custom Questions

In addition to the validated scales, five custom 7-point Likert-scale questions were analysed descriptively to provide supplementary insight into participants' experiences. In total the researcher presented 9 custom questions to the participants of the study, but only questions 2-6 allowed numerical answer from 1 to 7 on Likert's scale. These questions have not been academically validated and are therefore reported descriptively rather than inferentially, with two exceptions explored below.

Interaction intuitiveness (Markers_2): All groups found interaction intuitive, with similar means: control ($M = 5.75$, $SD = 1.36$), diegetic ($M = 6.00$, $SD = 1.21$), non-diegetic ($M = 5.91$, $SD = 1.51$). This consistency suggests that baseline usability did not differ across conditions, confirming that any differences (or lack thereof) in narrative immersion are not attributable to usability problems.

Sound interference of the main experience (Markers_3): Moderate scores across all groups: control (M = 4.75, SD = 1.76), diegetic (M = 4.58, SD = 1.93), non-diegetic (M = 4.18, SD = 1.83). The audio component of the experience was perceived similarly regardless of marker condition.

Marker recognition effectiveness (Markers_4): The non-diegetic group reported the highest recognition ease (M = 6.09, SD = 1.38), followed by diegetic (M = 5.58, SD = 1.38) and control (M = 4.75, SD = 2.09). This pattern suggests the explicit exclamation mark was most effective at communicating intractability, followed by the brightness shader, with no markers being least effective. However, a Kruskal-Wallis test found this difference was not statistically significant (p = .254).

Story ease (Markers_5): Nearly identical across groups: control (M = 4.58, SD = 1.88), diegetic (M = 4.67, SD = 1.23), non-diegetic (M = 4.64, SD = 1.86). This reinforces the NES Total finding that narrative comprehension was unaffected by marker type.

Markers and story importance (Markers_6): The non-diegetic group scored noticeably higher (M = 5.18, SD = 1.33) than control (M = 4.00, SD = 1.81) and diegetic (M = 3.75, SD = 1.82). This was the closest result to statistical significance in the entire dataset - a one-way ANOVA yielded F(2, 32) = 2.637, p = .087, with the diegetic–non-diegetic comparison at p = .117 in Tukey's post hoc test. While not reaching the .05 threshold, this trend suggests that non-diegetic markers may more effectively communicate narrative importance than diegetic markers, a finding worthy of further investigation with larger samples. It should be noted that applying inferential statistics to a single Likert item is methodologically limited, as individual items do not approximate continuous data in the same way that composite scores do. This result is therefore reported as an exploratory observation rather than a confirmable finding.

4.2.7 Summary of the descriptives

Table below sums up relevant descriptives (in order: median, mean standard deviation) of the inferential tests conducted in the section 4.2.4 for each of the subscales across conditions. While the individual test results have already been reported with their corresponding means and medians, this summary allows the reader to compare the central tendency and spread across all measures at a glance. The consistency of scores across conditions points towards descriptives falling within narrow ranges for each measure, reinforcing the null findings.

item	Control (n=12)	Diegetic (n=12)	Non-Diegetic (n=11)	Test Used
NES Total	μ = 5.15 $\tilde{x}$ = 5.29 σ = 1.15	μ = 5.01 $\tilde{x}$ = 5.00 σ = 0.704	μ = 5.11 $\tilde{x}$ = 5.42 σ = 1.06	ANOVA
NES Attention Focus	μ = 5.44 $\tilde{x}$ = 5.50 σ = 1.18	μ = 5.36 $\tilde{x}$ = 6.00 σ = 1.24	μ = 5.81 $\tilde{x}$ = 6.00 σ = 1.27	Kruskal-Wallis
NES Narrative Presence	μ = 5.19 $\tilde{x}$ = 5.33 σ = 1.33	μ = 4.86 $\tilde{x}$ = 5.00 σ = 1.13	μ = 5.24 $\tilde{x}$ = 5.00 σ = 1.26	ANOVA
NES Emotional Engagement	μ = 4.61 $\tilde{x}$ = 4.67 σ = 1.77	μ = 4.75 $\tilde{x}$ = 4.67 σ = 1.05	μ = 4.48 $\tilde{x}$ = 4.33 σ = 1.29	Kruskal-Wallis

IEQ-SF Involvement*	μ = 5.60 $\tilde{x}$ = 5.75 σ = 0.968	μ = 5.40 $\tilde{x}$ = 5.25 σ = 1.04	μ = 5.68 $\tilde{x}$ = 5.75 σ = 0.717	ANOVA
IEQ-SF Challenge*	μ = 2.89 $\tilde{x}$ = 2.83 σ = 1.41	μ = 2.61 $\tilde{x}$ = 3.00 σ = 1.31	μ = 2.64 $\tilde{x}$ = 2.67 σ = 0.767	ANOVA

** Exploratory; subscale reliability below acceptable threshold (α < .70)*
Table 4.5 - summary up relevant descriptives (in order: mean, median, standard deviation) of the inferential tests

Additionally, table 4.6 presents the corresponding descriptive statistics for the five custom Likert-scale questions reported in section 4.2.6. As noted, these items have not been academically validated and are reported as supplementary context. The one notable pattern visible in this summary is the Markers_6 item (markers and story importance), where the non-diegetic group scored noticeably higher than the other two conditions - the only measure in the entire dataset that approached statistical significance (p = .087).

item	Control (n=12)	Diegetic (n=12)	Non-Diegetic (n=11)
Markers_2 (Interaction intuitiveness)	μ = 5.75 σ = 1.36	μ = 6.00 σ = 1.21	μ = 5.91 σ = 1.51
Markers_3 (Sound enjoyment)	μ = 4.75 σ = 1.76	μ = 4.58 σ = 1.93	μ = 4.18 σ = 1.83
Markers_4 (Marker recognition)	μ = 4.75 σ = 2.09	μ = 5.58 σ = 1.38	μ = 6.09 σ = 1.38
Markers_5 (Story Ease)	μ = 4.58 σ = 1.88	μ = 4.67 σ = 1.23	μ = 4.64 σ = 1.86
Markers_6 (Markers and story importance)	μ = 4.00 σ = 1.81	μ = 3.75 σ = 1.82	μ = 5.18 σ = 1.33

Table 4.6 - Summary of descriptives (in order: mean, standard deviation) of the custom Markers questions

4.3 QUALITATIVE DATA ANALYSIS

To complement the quantitative findings, qualitative data was gathered from three sources: the open-ended question in the questionnaire ("What did you like about the design of the markers present in your VR experience?"), verbal comments made by participants after the gameplay experience, and the researcher's observations during the testing sessions. Following Braun and Clarke's (2006) thematic analysis approach, the researcher familiarised herself with the data, generated initial codes from recurring patterns across participant responses and observations, and refined these into three themes: marker perception and clarity across conditions, sound design as an influence on immersion, and interaction behaviour with non-interactable objects.

4.3.1 Marker clarity: non-diegetic markers perceived as clearest across conditions

One of the open-answer questions asked to the participants was "What did you like about the design of the markers present in your VR experience?" that was asked to get more in-depth, recorded response to analyse in the following section.

4.3.1.1 Control group

Control group participants had no distinctive markers in their VR experience, therefore only 5 out of 12 participants of this condition left a comment in the question Markers_9, that can be useful for the analysis.

Among those who did respond, comments focused on the overall aesthetic of the environment rather than markers specifically. One participant noted that the design felt connected to the story and child-like feeling of the environment - "the design felt connected to the story told by a child. (I) felt as it was made by the same child (I was playing)". Another participant appreciated the proportions and textures of the virtual space - "I liked the textures of the space and proportions, objects were a little bit bigger than in the real world, but it matched the story and the vibe it was giving off.".

Interestingly, one participant mentioned that "the outline made it intuitive about which toys are to be interacted with," suggesting they recognised the slight visual distinction in the interactive objects, compared to the 3D scanned environment. Those differences were present in the original game that was developed and presented in the graduation exhibition in 2024, and they were noted in the Methodology chapter. This observation could be interpreted through Gibson's (1979) affordance theory, discussed in section 2.2.1: the toys' physical properties - their shape, scale, and visual distinction from the scanned furniture may have communicated interaction possibilities without any designed signifier; and Steffen et al. (2019) whose framework validated that VR objects carry inherent affordances from their physical-world counterparts. In other words, the objects themselves may have carried sufficient natural affordances to guide player behaviour, which could partly explain why the control group's narrative immersion scores were comparable to the marker conditions. This observation also raises the possibility that the control condition was not entirely free of visual affordance cues, which may have reduced the effective perceptual difference between conditions.

4.3.1.2 Diegetic markers

One out of twelve participants didn't report their answer to the Markers_9 question. Responses in the diegetic condition were notably mixed regarding clarity. Several participants found the brightness markers effective and immersion-preserving, describing them as something that "clearly stood out, but didn't break immersion" and as "not too obvious, but intuitive."

One participant highlighted the inclusivity potential, suggesting that older generations would benefit from this feature. - "I like that for me it was more obvious what to reach out to, I also believe that older generation would benefit from this feature to make it more reachable for them (...)".

However, other participants expressed difficulty identifying what the markers were: one noted that "the colours were very nice of the markers, it was not clear to me what I could grab and what I couldn't. For example I wanted to grab the chair but didn't know I could not. Maybe it' s because I almost have no experience in gaming. (...)". Another remarked that " the fact that while the interactible items were quite clear, it was a bit difficult to tell what the exact markers were that identified the interactable items ".

This mixed reception could be understood through Norman's (1988) distinction between actual affordances and perceived affordances, discussed in section 2.2.1 of the literature review. The brightness marker created a signifier - a visual cue intended to communicate interactability, but not all participants perceived it as such. For those with gaming experience, the subtle brightness shift may have been sufficient to signal "this object is interactive." For participants with little or no gaming background, the same visual cue may not have registered as a deliberate signifier, instead blending into the overall aesthetic of the environment. This pattern illustrates Norman's point that an affordance goes unused if users do not perceive it - the interaction possibility

existed, but the diegetic marker did not reliably communicate it to all players. It also suggests that diegetic markers, functioning as motivated signs that rely on natural perceptual properties rather than learned conventions (as discussed in section 2.2.2), may carry an inherent trade-off: they preserve immersion for those who perceive them, but risk being overlooked by those who lack the perceptual framework to interpret them as intentional signals.

4.3.1.3 Non-Diegetic markers

One participant out of eleven didn't answer the Markers_9 question. The non-diegetic condition received the most consistently positive feedback regarding clarity. Participants described the exclamation mark markers as "very visible," " clear to understand", "easy to spot from a distance (...)", and effective at allowing players to "easily differentiate them from the non-interactable items."

Notably, several participants indicated that the markers did not feel disruptive - one described them as "not necessarily distracting, but if I needed guidance they were there," and another mentioned not paying much attention to their design but appreciating their "crayon-y sketch-y look." This last observation is particularly relevant, as it suggests the stylistic integration of the non-diegetic markers into the child's bedroom aesthetic (discussed in the methodology chapter) may have softened their conventionally game-like appearance.

It connects to the semiotic framework outlined in the literature review: the non-diegetic exclamation marks function as conventional signs (Saussure, 1917) with clear learned associations, while the diegetic brightness functions as a motivated sign that communicates through natural perceptual properties but carries inherent ambiguity (Fagerholt & Lorentzon, 2009).

4.3.2 Sound design as potential confounding variable across conditions

Having completed the 7-10min play in the VR headset participants, among other control questions specified in the methodology chapter, were encouraged to express their thoughts after the experience verbally. Those conversations revealed a theme of layered sounds from the "Echoes of Abandonment" could have directly influenced the immersion of participants.

During post-gameplay conversations, several participants across all conditions commented on the sound design of "Echoes of Abandonment," suggesting an unintended tension in the game's core mechanic. As described in the methodology, the game's central interaction involves picking up toys that each produce musical sounds, which together compose a musical piece that gradually mutes the parents' argument heard behind the door.

Several participants reported that while they found the game engaging, the layered sounds - particularly the xylophone, which plays for a longer duration (1min30sec) - became a source of annoyance rather than immersion. One participant with a professional background in sound design specifically noted that the xylophone's sustained playback could lower immersion, as players may become irritated rather than engaged by the sound. He expressed "the longer the xylophone was playing the lower my sense of immersion became". More broadly, multiple participants indicated that the musical composition created by toy interaction actually conflicted with their desire to listen to and understand the parents' argument, which formed the narrative core of the experience.

This possibility connects to Brooks' (2003) argument that complex visual design is unnecessary when audiences are already deeply engaged in narrative events. It could also be understood through Murray's (1997) observation about the fragility of immersive experiences - if

immersion in "Echoes of Abandonment" was primarily sustained through the emotional weight of the audio narrative, visual UI elements may not have been powerful enough to break that engagement, regardless of their diegetic classification. This finding complicates the picture presented by Pillai and Verma (2019), who suggested that technical and narrative elements support rather than compete with each other. In 'Echoes of Abandonment,' the sound mechanic appears to have created friction rather than synergy with the narrative content. This topic will be discussed more in depth in the Discussion chapter.

In other words, the game mechanic designed to enhance engagement (creating music) may have been perceived by some participants as competing with the narrative content.

4.3.3 Exploratory behaviour differences: informal observations on static object interaction

From the observations of the researcher while participants were taking part in the study, it was noticed that in some conditions participants attempted to interact with non-interactable objects (such as the chair, bed, or door) more often than in others. The researcher observed that participants in the control group and non-diegetic condition appeared to attempt interaction with static objects more frequently, while participants in the diegetic condition appeared to do so less often.

This observation should be interpreted with caution, as it is based on the researcher's informal impressions during testing rather than systematic behavioural recording. No structured observation protocol was used, and no quantitative count of interaction attempts was maintained – the observations were a by–product of the methodology. Future research could address this by implementing interaction logging within the game engine to capture precise data on player behaviour.

4.3.4 Summary of the qualitative findings

The qualitative analysis revealed three key insights that complement the quantitative results. First, based on the answers in the questionnaire, responses from the non-diegetic group suggest that participants tended to perceive the exclamation mark markers as clearer, though the small sample and open-ended nature of the question mean this pattern should be interpreted as indicative rather than conclusive. Diegetic markers received more mixed responses, some participants appreciated their subtlety and immersion preservation, while others, particularly those with less gaming experience, found them insufficiently clear as signifiers.

Secondly, the sound design of "Echoes of Abandonment" may have acted as a confounding variable by creating frustration across all conditions, potentially dampening narrative immersion scores uniformly.

Thirdly, informal observations suggested possible differences in exploratory behaviour across conditions, though this finding requires systematic investigation to confirm. Together, these qualitative insights help contextualise the null quantitative findings by revealing experiential nuances that standardised scales could not capture.

4.4 HYPOTHESIS EVALUATION

According to the ANOVA and Kruskal-Wallis statistical tests the type of marker didn't significantly affect participants' narrative engagement. Following this conclusion the researcher

shall reject Hypothesis 1 - "The diegetic markers will support the narration immersion.", and Hypothesis 2 - "The non-diegetic markers will noticeably disturb the narration immersion of the VR game players". Therefore Hypothesis 0 cannot be rejected, meaning there is no influence of marker types on player's immersion into narration.

4.4.1 Interpretation of null findings

The absence of statistically significant differences across all three marker conditions, with effect sizes consistently below $\eta^2 = .02$, calls for elaboration. Rather than simply concluding that markers do not matter, it is important to examine what factors may have contributed to this outcome. Possible explanations are explored in the following section.

4.4.1.1 Sample size and effect detection

With 11–12 participants per condition, this study's sample is consistent with comparable VR research conducted within academic settings and similar practical constraints (Köhle et al., 2021, n = 37; Saling et al., 2021, n = 13). Nonetheless, smaller group sizes inherently limit the ability to detect small effects, and the negligible effect sizes observed here (all $\eta^2 < .02$), leave open the question of whether the null findings reflect a genuine absence of effect or whether a subtle difference exists that would only emerge with substantially larger samples. This ambiguity is common in exploratory VR research and does not diminish the value of the findings, but it does mean the results should be understood as indicative rather than definitive.

4.5 LIMITATIONS

This section consolidates the methodological limitations identified throughout the data collection and analysis process.

4.5.1 Scale differences

It should be noted that the IEQ-SF was administered using a 7-point Likert scale in this study, whereas the validated administration format specifies a 5-point scale (Cutting et al., 2025). This deviation may have contributed to the reliability issues observed. The 7-point scale has been used to ensure consistency across the questionnaire. Additionally, the IEQ-SF was developed and validated using data from screen-based gaming contexts, where immersion manifests differently than in VR. In particular, the Real-World Dissociation items ask about awareness of real-world surroundings - a construct that may behave fundamentally differently when participants are wearing a VR headset that physically blocks the real world, compared to playing on a screen where the real world remains visible. The Challenge items also showed weak performance, which aligns with the narrative exploration nature of "Echoes of Abandonment" - a game designed for emotional storytelling rather than skill-based challenge.

4.5.2 Participants

The sample size of 35 participants, while consistent with comparable VR studies conducted within academic settings (Köhle et al., 2021, n = 37; Saling et al., 2021, n = 13), limits the statistical power available to detect small effects. With 11–12 participants per condition, the study was able to identify large or medium effects, but subtle differences between marker conditions, might not have reached statistical significance.

This constraint reflects the practical realities of conducting a VR experiment within a master's thesis timeline: each session required individual supervision, dedicated physical space, and specialised hardware, making large-scale recruitment unfeasible within the four-week data collection period. To mitigate this limitation, the researcher ensured methodological rigour in other areas - consistent procedural conditions across all sessions, validated measurement instruments, and systematic assumption testing. so that the data collected, while limited in volume, would be as reliable as possible. The negligible effect sizes observed (all $\eta^2 < .02$) suggest that marker type may genuinely have minimal influence on narrative immersion in this context, but replication with larger samples would be needed to confirm this interpretation.

Secondly, the relatively high proportion of participants from game-related backgrounds (34.3%) means that over a third of the sample likely had above-average familiarity with game UI conventions, including marker systems. This familiarity may have influenced how these participants perceived and interpreted the diegetic and non-diegetic markers used in this study. The distribution of professional backgrounds across conditions was examined to assess whether game-related participants were evenly spread or clustered in particular groups.

4.5.2.1 Missing answer

One participant (participant 24, non-diegetic condition) did not respond to IEQ-SF item 4 ("I enjoyed playing the game"), which belongs to the Involvement subscale. Given that the IEQ-SF already demonstrated poor overall reliability in this study ($\alpha = .487$), and the Involvement subscale was retained only for exploratory analysis, this single missing data point is unlikely to have meaningfully influenced the results. The analysis proceeded with this participant's remaining responses included, as excluding an entire participant's data over one missed item would have further reduced the already small non-diegetic group from 11 to 10 participants.

It should be noted that the missing response meant the Involvement subscale ANOVA was conducted with slightly different degrees of freedom ($F(2, 31)$ rather than $F(2, 32)$) compared to other analyses. This is reflected in the reported results. Had the missing value been addressed through imputation — for instance, substituting the group or item mean — the results would likely have remained unchanged given the negligible effect sizes observed across all measures. However, imputation was not applied, as introducing an estimated value into an already unreliable subscale could have introduced additional measurement noise without meaningful benefit.

4.5.3 Audio-primary narrative delivery

In "Echoes of Abandonment" the story is delivered almost entirely through audio - the parents' argument heard behind the door - rather than through the visual interaction with the marked toys. If participants were primarily immersed through listening to the unfolding dialogue, the visual markers on interactable objects may have been peripheral to their narrative experience. In other words, the markers existed on a sensory channel (visual) that was secondary to the channel through which the narrative was being conveyed (auditory).

This interpretation is supported by the qualitative findings of this study. As discussed in section 4.3.2, several participants reported that the toy sounds competed with, rather than complemented, the parents' argument - suggesting that their narrative attention was directed toward the audio content. If this was the case, it could explain why neither marker type

disrupted or enhanced narrative immersion: the markers may simply not have been salient enough relative to the auditory narrative to register as either a help or a hindrance.

4.5.4 Collection of data in various locations

While the researcher ensured consistent procedural conditions across all testing locations, the specific location was not systematically recorded per participant, preventing post hoc analysis of potential location effects. The consistent use of identical hardware, software, and questionnaire administration across all sessions was implemented to minimize any environmental influence.

4.6 ETHICAL CONCERNS

The ethical concerns outlined in the methodology chapter (section 3.8) were implemented throughout data collection. This section reports on how ethical considerations manifested during the actual study.

Regarding participant wellbeing, three participants reported mild headache or dizziness after removing the VR headset. In each case, the researcher immediately offered water and seating, and remained with the participant until they confirmed they felt comfortable proceeding with the questionnaire. No participant chose to withdraw from the study as a result of physical discomfort. These symptoms are consistent with commonly reported effects of VR use documented in the literature (Chattha et al., 2020) and were anticipated in the ethical considerations of the methodology chapter, where rest periods between the VR experience and questionnaire completion were built into the procedure for this reason.

No participants reported emotional distress related to the narrative content of "Echoes of Abandonment," despite the potentially sensitive theme of parents arguing. This may be partly attributable to the informed consent process, which described the nature of the audio content in advance, giving participants the opportunity to decline before exposure. However, it should be noted that social desirability effects, particularly given the researcher's presence and role as the game's creator, may have discouraged participants from expressing discomfort.

All participant data was collected and stored anonymously, as specified in the methodology chapter. No personally identifiable information was linked to questionnaire responses at any point during data collection or analysis. The data was stored securely on the Qualtrics platform and will be deleted within 30 days of final thesis submission and grading, in accordance with the commitment made in the consent form.

5 DISCUSSION OF DATA AND RESULTS

The purpose of this study is to investigate the research question "*What is the influence of different markers design in VR games on the player's ability to become immersed in narrative content among young adults living in the Netherlands?*". While the quantitative analysis returned null findings across all three marker conditions, the qualitative data and observational notes revealed meaningful experiential differences between participants. Rather than treating the null result as the end of the inquiry, the following sections examine what the data, taken together, suggest about the relationship between marker design and narrative immersion in VR.

5.1 INTERPRETATION OF THE NULL HYPOTHESIS

The absence of significant differences between the diegetic, non-diegetic, and control conditions across all narrative engagement and immersion measures invites careful interpretation. Rather than indicating that marker design is irrelevant to narrative immersion in VR, the null result is best understood by positioning the present study within an emerging pattern in VR UI research.

The most direct comparison is to Saling et al. (2021) and Köhle et al. (2021), both of which examined diegetic versus non-diegetic UI in VR shooter contexts. Saling et al. could not establish a strong relationship between UI design and spatial presence, while Köhle et al. concluded that different UI approaches were suited to different contexts rather than one being inherently superior. The present study extends this pattern to narrative immersion, a construct that the literature review hypothesised might respond differently than spatial presence due to its reliance on sustained cognitive and emotional engagement with story. The data did not bear out this distinction. Across three independent VR UI studies - two on spatial presence, one on narrative immersion - no significant effect of UI type has been observed. Taken together, these convergent null results begin to **suggest that the diegetic/non-diegetic distinction, as theorised by Fagerholt and Lorentzon (2009) for screen-based games, may not transfer cleanly to VR in the way the framework's continued use in the field implies.**

This does not, however, lead to the conclusion that marker design has no influence on narrative immersion. Three considerations restrain that interpretation. First, the negligible effect sizes observed in this study (all $\eta^2 < .02$) are consistent with both a genuine absence of effect and a small effect that the present sample lacked the statistical power to detect; this ambiguity has already been acknowledged in section 4.4.1.1 and is common in exploratory VR research. Second, the qualitative findings reported in section 4.3 revealed perceptual differences between conditions - particularly around marker clarity and recognition - that the validated quantitative scales did not capture. The null quantitative result therefore coexists with meaningful experiential variation, suggesting that immersion as operationalised by the NES and IEQ-SF may not be the construct most sensitive to marker manipulation. Third, several contextual factors specific to *Echoes of Abandonment* - the audio-primary narrative delivery, the high proportion of game-experienced participants, and the natural affordances of the toy objects themselves - may have moderated the effect in ways that the experimental design could not isolate. These factors are examined in turn in sections 5.1.2 through 5.1.4.

In direct response to the research question - *"What is the influence of different markers design in VR games on the player's ability to become immersed in narrative content among young adults living in the Netherlands?"* - the present study finds no measurable influence of marker type on narrative immersion under the conditions tested. This answer should be read as conditional

rather than absolute: future research could strengthen the basis for this conclusion by replicating the study with a larger sample stratified by gaming experience, by testing the same marker conditions in a visually-driven rather than audio-driven narrative VR experience, and by complementing validated immersion scales with measures sensitive to moment-to-moment perceptual clarity. The remainder of this chapter explores why this null result emerged and what it suggests, both methodologically and theoretically, for the design of UI in narrative VR.

5.1.1 Theoretical interpretation: Communicating affordance

5.1.2 Theoretical interpretation: UI type as context-dependent

A broader theoretical explanation can be drawn from Köhle et al. (2021), whose study of diegetic versus non-diegetic UI in VR shooter games concluded that different UI approaches are suited to different contexts, rather than one type being inherently superior. The present study's null findings could be interpreted as extending this conclusion to narrative VR experiences: perhaps marker design does not strongly influence narrative engagement when the narrative is conveyed through a channel that operates independently of the visual markers.

This interpretation would suggest that the relationship between UI design and narrative immersion may be mediated by the degree to which the UI elements and the narrative content share the same sensory modality. In a VR game where the story is told visually, through environmental storytelling, text, or visual events, markers that disrupt the visual coherence of the game world might have a measurable impact on narrative immersion. In "Echoes of Abandonment," where the story unfolds auditorily, the visual markers may have occupied a different attentional space entirely.

If this interpretation holds, it would have practical implications for VR designers: the choice between diegetic and non-diegetic markers may matter most when the narrative and the UI compete for the same sensory channel, and matter less when they operate on separate channels. However, this proposition remains speculative based on a single study with a small sample, and would require targeted investigation - for instance, comparing marker effects in audio-driven versus visually-driven VR narratives - to be substantiated.

It is also worth noting that the null quantitative findings do not mean that participants experienced all three conditions identically. The qualitative data revealed meaningful perceptual differences: non-diegetic markers were described as clearer and more immediately recognisable, while diegetic markers received mixed responses regarding their communicative effectiveness. These experiential differences may exist on dimensions that the NES and IEQ-SF were not designed to capture - such as moment-to-moment confidence in knowing what to do, or the subjective feeling of being "guided" versus "discovering", rather than being absent altogether.

5.2 MEASUREMENT VALIDITY AND THE NEED FOR VR-NATIVE IMMERSION SCALES

Prior to data collection, this thesis evaluated the validated immersion and narrative engagement scales available in HCI and game studies research (Bareišytė et al., 2024), and selected the IEQ-SF and NES as the most appropriate instruments to answer the research question (see section 3.1). The reliability results obtained in the present study (NES α = .869; IEQ-SF α = .487) are, in

themselves, a substantive finding rather than a methodological footnote. The two scales, both well-established in screen-based gaming research and both administered to the same participants in identical conditions, produced markedly different outcomes. The NES held up across three of four subscales; the IEQ-SF fell well below the .70 acceptability threshold, with Real-World Dissociation reaching only α = .439. This divergence points to a structural difference in how each scale's underlying construct survives the transition from screen-based gaming to VR, what a difference that, once examined, has implications well beyond the present study.

The NES measures narrative engagement through items focused on attention to story, emotional response, and the experience of being absorbed in a fictional world. The NES measures narrative engagement through items focused on attention to story, emotional response, and the experience of being absorbed in a fictional world. The strong reliability obtained in this study (α = .869) is consistent with the scale's psychometric performance in screen-based contexts (Busselle & Bilandzic, 2009), suggesting that the underlying constructs survived the transition to VR delivery intact.

The IEQ-SF's poor performance, by contrast, points to a more specific issue. Its Real-World Dissociation subscale (α = .439) measures the degree to which players become unaware of their physical surroundings - a construct that varies meaningfully in screen-based gaming, where awareness of one's living room is a spectrum, but that the VR headset eliminates by removing the real-world visual field. This may explain why Real-World Dissociation showed the weakest reliability of any subscale in this study: the items ask about a state the equipment has already produced, leaving little variance for the scale to capture.

This pattern of reliability concerns is consistent with the broader methodological context in which the IEQ-SF itself was developed. Cutting et al. (2025), motivated the creation of the shortened form of the original IEQ (Jennett et al., 2008), developed using methods that no longer met current psychometric standards, positioned the IEQ-SF as a more rigorous and practically usable alternative. In their validation work, they concluded that the IEQ-SF was "a good substitute for the full IEQ in both high and low immersion situations." The present study's results extend this conversation in an important direction: while the IEQ-SF appears to function as Cutting et al. intended within the screen-based gaming contexts where it was validated, its reliability is notably lower when transferred to a VR narrative context, with Real-World Dissociation showing the weakest reliability of any subscale measured (α = .439). This suggests that the substitutability claim, while well-supported within its original validation scope, may not extend cleanly to immersive media where the medium itself reshapes the constructs the scale is designed to capture.

The reliability divergence observed in this study points to a broader structural issue in VR immersion research. Both the NES and the IEQ-SF, like the majority of validated immersion and engagement scales available to researchers today, were developed and psychometrically validated using data from screen-based contexts: the NES for narrative film, television, and reading (Busselle & Bilandzic, 2009), and the IEQ-SF for screen-based digital games (Cutting et al., 2025). This was a reasonable starting point, therefore researchers studying VR immersion adopted the best available instruments and adapted them to the new context. The present study illustrates the limits of that adaptive approach. The constructs themselves, particularly those built around the player's relationship to the real-world environment, were operationalised under assumptions that no longer hold when the medium physically removes that environment from the participant's perceptual field. As VR matures as both a consumer technology and a

research medium, the field is reaching a point where adapted instruments are insufficient and purpose-built ones become necessary. Further directions for the future researchers will be discussed in the 7.1 chapter of this thesis.

Two implications follow for how the present study's results should be interpreted. The NES findings, supported by strong reliability across the relevant subscales, can be read with full interpretive weight: the null result on narrative engagement reflects a genuine pattern in the data rather than a measurement artefact. The IEQ-SF findings, by contrast, should be read with the unreliability explicitly in view; the null results on its subscales are best understood as consistent with a genuine null effect but not the strong evidence of one, since the instrument's capacity to detect group differences in this context is questionable from the outset. This is precisely the kind of distinction that VR-native instruments would help future researchers draw with greater confidence and, in turn, the kind of nuanced interpretive position that the present study's reliability findings make available to the field.

5.3 THE NATURAL AFFORDANCES FINDING FROM THE CONTROL GROUP

One of the more theoretically consequential findings of this study emerged from a condition that was designed to produce natural affordances. The control group, by experimental design, received no marker treatment, as participants in this condition played "Echoes of Abandonment" with neither the diegetic brightness shader nor the non-diegetic exclamation mark applied to interactable toys. The expectation underlying this design was that the absence of designed signifiers would produce lower marker recognition scores and, therefore, establish a baseline against which the two marker conditions could be compared. The data did not support this expectation. Control group participants reported interaction intuitiveness scores ($M = 5.75$) statistically indistinguishable from those of the diegetic ($M = 6.00$) and non-diegetic ($M = 5.91$) conditions, and their narrative engagement scores fell within the same narrow range as the two marker conditions.

While most control group participants did not comment on markers in their open-ended responses - unsurprisingly, given that no designed markers were present - one participant explicitly noted that "the outline made it intuitive about which toys are to be interacted with", referring to the slight visual distinction between the interactable toys and the 3D-scanned environment of "Echoes of Abandonment". This visual distinction, documented in the methodology chapter as a residual feature of the original game's production, was not part of the experimental manipulation. It appears to have functioned as a signifier in the Normanian sense - a perceivable property that communicated interactability to at least some participants. Other participants in this condition described the toys' aesthetic integration with the child-storyteller framing of the environment without referencing markers, suggesting that the toys' visual properties were salient enough to feel coherent within the world rather than absent of communicative function.

This pattern invites a reinterpretation of what the control condition actually tested. Gibson's (1979) ecological framing of affordances holds that interaction possibilities are not added to objects through design but exist in the relationship between an object's physical properties and a perceiver's capabilities. From this perspective, the toys in "Echoes of Abandonment" - hand-sculpted in clay, scaled to a child's perception, visually differentiated from the scanned room they sit within - already carried affordances for pick-up interaction independent of any marker

applied to them. Steffen et al.'s (2019) framework of VR-specific affordances reinforces this reading: their empirical validation work established that VR objects inherit affordances from their physical-world counterparts, meaning that a virtual toy designed to look like a toy already signals "this can be picked up" to a perceiver who has ever picked up a toy before. The control group, on this account, was not a no-signifier condition. It was a condition in which the only signifiers available were the natural ones the objects carried by virtue of being recognisable toys in a recognisable room.

If this reading holds, the three experimental conditions can be understood as differing in the type and noticeability of the affordance signification available to participants rather than in its presence or absence. The diegetic brightness functioned as a motivated sign layered on top of natural perceptual properties; the non-diegetic exclamation mark functioned as a conventional sign requiring learned association, also layered on top of those same natural properties; and the control condition relied on the natural properties alone. Framed this way, all three conditions provided participants with some perceivable basis for distinguishing interactable from non-interactable objects, and the comparison between conditions becomes one between three forms of affordance communication rather than between marker presence and marker absence.

This framing does not invalidate the null finding, but it does offer a plausible contributing explanation for it alongside the other moderating factors examined in this chapter. Rather than indicating that marker design does not influence narrative immersion, it may indicate that narrative immersion is robust across multiple effective forms of affordance communication, provided that affordances are communicated at all. The diegetic and non-diegetic markers were not competing against an absence of communication but against the inherent semiotic richness of the objects they were applied to. For a marker condition to produce a measurable difference, it would have to either substantially amplify what the object already communicates or actively disrupt the player's perception of it - a higher bar than the experimental design implicitly assumed.

The methodological lesson for future VR UI research follows directly. A true no-signifier control in a VR environment populated by recognisable objects may be difficult or even impossible to construct, since the recognisability of the objects is itself a signifier. Researchers seeking to isolate the effect of designed markers may need to either work with abstract or unfamiliar virtual objects whose affordances cannot be inferred from real-world counterparts, or to reconceive the comparison entirely - shifting from "marker versus no marker" to "marker type A versus marker type B versus natural affordance only", with the latter treated not as a baseline but as a third experimental condition in its own right. This reconception is consistent with the direction Köhle et al. (2021) gestures toward - that UI design questions in VR are better posed as comparisons between effective design strategies than as comparisons between presence and absence of UI.

5.4 "THE INTERACTION MARKERS HELPED ME UNDERSTAND WHAT WAS IMPORTANT TO THE STORY" SIGNIFICANT SCORE

e supplementary context for interpreting the validated scale results. One item within that set, however, produced a pattern that warrants closer attention. Markers_6 asked participants to rate the statement "The interaction markers helped me understand what was important to the story" on a seven-point Likert scale. The non-diegetic group scored noticeably higher on this

item (M = 5.18, SD = 1.33) than either the diegetic (M = 3.75, SD = 1.82) or the control condition (M = 4.00, SD = 1.81). A one-way ANOVA yielded F(2, 32) = 2.637, p = .087, and Tukey's post hoc comparison between the diegetic and non-diegetic groups returned p = .117. This was the closest result to statistical significance in the entire dataset, and the only one across both validated scales and custom items where the non-diegetic condition diverged meaningfully from the other two.

The methodological caveat applied to this result in the results chapter remains in force here. Applying inferential statistics to a single Likert item is limited in ways that composite scale scores are not - individual items do not approximate continuous data, and a single item cannot capture the construct it gestures at with the reliability that a multi-item scale can. Neither the ANOVA nor the post hoc comparison crossed the conventional .05 threshold, and the difference in means, while visible, rests on group sizes of eleven and twelve. For these reasons, the present discussion treats the result as an exploratory observation worth interpreting rather than as a finding to be defended.

With those caveats in view, the direction of the result is itself theoretically interesting because it sits orthogonally to the hypotheses this thesis set out to test. Hypothesis 2 predicted that non-diegetic markers would "noticeably disturb the narration immersion" of VR players, on the grounds that conventional signs operating outside the diegesis would remind players of the game interface and break their engagement with the story. The Markers_6 trend does not contradict that prediction directly, since narrative engagement as measured by the NES was not lower in the non-diegetic condition. What the trend suggests instead is that non-diegetic markers may have done something the hypotheses did not anticipate them doing: rather than disturbing narrative engagement, they may have actively aided participants in identifying which elements of the environment carried narrative weight.

This reading can be situated within the theoretical frameworks established in the literature review. Norman's (2013) account of signifiers describes them as design elements that communicate not only the existence of an affordance but also its relevance - a signifier tells the user not just that an action is possible, but that the system considers it worth doing. The non-diegetic exclamation mark, as a conventional sign with strong learned associations in gaming culture (Saussure, 1917; Fagerholt & Lorentzon, 2009), carries this relevance signal more explicitly than a motivated sign like brightness does. A glowing toy says "I am different from the surrounding objects"; an exclamation mark says "look here, this matters". In a narrative VR experience where the player must work out which interactive elements carry story significance, the second message may be more useful than the first. The qualitative data reported in section 4.3.1.3 lends additional support to this interpretation, since non-diegetic markers were repeatedly described as "easy to spot from a distance" and effective at allowing players to "easily differentiate them from the non-interactable items".

If this interpretation holds, it points to a more nuanced relationship between marker type and narrative immersion than the diegetic/non-diegetic framework alone would predict. The Fagerholt and Lorentzon (2009) framework treats diegetic UI as more immersion-preserving by default, on the assumption that anything stepping outside the story world risks pulling the player out of it. The Markers_6 trend complicates that assumption by suggesting that what the player needs to remain narratively engaged is not necessarily an absence of interface but a clear sense of where the story's interactive weight lies. Non-diegetic markers, despite their conventional and "game-like" character, may serve narrative immersion by reducing the

cognitive cost of identifying narratively relevant interaction points, leaving more of the player's attention available for the story itself.

This interpretation should be read as a hypothesis generated by the present study rather than as a finding established by it. A targeted replication with a larger sample, an inferentially valid composite measure of "narrative salience communication", and ideally a within-subjects design exposing each participant to multiple marker types within the same narrative would be needed to determine whether the trend observed here reflects a real effect. What can be said now, on the basis of the present data, is that the relationship between non-diegetic markers and narrative immersion is unlikely to be the straightforwardly disruptive one Hypothesis 2 proposed, and that the question of how different marker types support narrative comprehension - as distinct from how they affect narrative engagement - is one the field has not yet asked clearly and that this study's data suggests is worth asking.

5.5 SOUND DESIGN AFFECTING THE RESULTS

Among the qualitative themes identified in the analysis chapter, the most consistent across conditions was a tension between the game's central audio mechanic as narrative content delivery method and its layering sound. As reported in section 4.3.2, several participants commented during post-gameplay conversations that the layered sounds produced by interacting with toys - particularly the xylophone, with its sustained one-minute-thirty-second playback - became a source of annoyance rather than engagement. One participant with a background in sound design articulated this directly: "the longer the xylophone was playing the lower my sense of immersion became". More broadly, multiple participants across all three conditions indicated that the musical composition created by toy interaction competed with, rather than complemented, their attention to the parents' argument that formed the narrative core of "Echoes of Abandonment". This pattern emerged without prompting from the researcher and was not anticipated by the experimental design, which treated sound as a constant across conditions rather than as a variable likely to influence the dependent measure.

This pattern emerged unanticipated and is worth examining, because the existing literature on sound design in immersive media suggests it should not have been surprising. Sound is widely understood by both academic researchers and industry practitioners to be a primary, not secondary, channel of immersion in VR. Murray's (1997) account of immersive media positions multi-sensory coherence as foundational to the immersive experience, and Brooks (2003) similarly argues that complex visual design becomes secondary when audiences are deeply engaged with audio narrative content. Within games and VR specifically, empirical work has established that auditory cues meaningfully contribute to presence and immersion: Hendrix and Barfield (1996) demonstrated that the addition of spatialized sound significantly increased participants' sense of presence in stereoscopic virtual environments, and subsequent research has consistently identified auditory factors - spatial properties, background soundscape, cross-modal consistency, and audio content quality - as core contributors to the immersive experience (Larsson et al., 2010; Brinkman et al., 2015). Industry accounts converge on the same point: in audio design for narrative VR experiences, sound is described as "half of the experience", and lacking variation or producing fatiguing repetition is identified as a primary mechanism by which immersion is lost (Chang, 2017). The unexpected finding of the present study is therefore unexpected only in the narrow sense that the experimental design did not predict it; in the

broader context of audio-immersion research, the pattern observed in participants is exactly what that literature would lead a researcher to expect.

This reading has direct implications for how the null quantitative findings should be interpreted. If sound design produced friction with narrative engagement across all three conditions equally, then any influence of marker design on narrative immersion would have had to register on top of an already-degraded baseline. In statistical terms, this is a candidate confounding variable that operated uniformly across the experimental groups, dampening overall immersion scores and reducing the variance that the marker manipulation could have produced. The NES mean scores reported in section 4.2.4.1 - ranging tightly from 5.01 to 5.15 across conditions, well within the upper-middle range of the seven-point scale but not approaching ceiling - are consistent with this account: participants were moderately, not strongly, immersed, and the moderation appears to have been consistent across groups.

The finding also enters into conversation with the theoretical model proposed by Pillai and Verma (2019), discussed in section 2.3.2 of the literature review. Their account of narrative immersion in VR cinema distinguishes between technical immersion (sustained through attentional cues and environmental fidelity) and narrative immersion (arising from story, character, and viewer integration), and argues that these two layers can support and reinforce each other to strengthen the overall immersive experience. The present study's qualitative data does not contradict that claim, but it does suggest a condition under which their model may not hold: when the technical or mechanical layer of a VR experience produces sound that competes for the same sensory channel as the narrative content. In "Echoes of Abandonment", the sound-producing toy mechanic and the audio-delivered narrative occupied the same auditory bandwidth, and rather than reinforcing each other in the manner Pillai and Verma describe, they appear to have created interference. This suggests that the supportive relationship between technical and narrative immersion that their model proposes may be conditional on the two layers operating either on different sensory channels or on the same channel without competing for attentional focus. This is a refinement of their model rather than a refutation, but it is the kind of refinement that the field of narrative VR research has not yet articulated clearly.

A further factor that may have shaped the strength of some participants' reactions to the layered sound design is the possible presence of auditory sensitivity within the sample. The questionnaire did not collect data on participants' neurodivergence, and so no claim can be made about the actual composition of the sample on this dimension. However, recent Dutch population data indicates that approximately 3% of people aged four and above self-report having autism spectrum disorder (CBS, 2025), and 1–2.5% of Dutch adults aged 18–44 meet diagnostic criteria for ADHD (de Graaf et al., 2014); and both figures may underestimate true prevalence, as adult cases often go undiagnosed. A sample of thirty-five participants drawn from the general young-adult population in the Netherlands could plausibly have included individuals who experience repetitive or layered sounds more intensely than the neurotypical norm. Auditory hyper-reactivity is a documented feature of both autism spectrum and ADHD presentations, with empirical work showing that neurodivergent individuals report substantially higher rates of sound tolerance problems and auditory distractibility than neurotypical controls (Kuiper et al., 2024). The mechanism described in this literature - reduced filtering of competing auditory input, leading to sensory overload when multiple sounds occur simultaneously - maps closely onto the experience reported by several participants in the present study, who described the musical composition created by toy interaction as conflicting with their desire to listen to and understand the parents' argument that formed the narrative

core of the experience. Without diagnostic data this can only be offered as a plausible contributing factor rather than a confirmed one, but it is plausible enough that it warrants attention in any replication of this study and in the design of future narrative VR experiences intended for general audiences.

Two implications follow from this analysis. The first concerns the present study's findings: the null result on marker conditions should be interpreted in the knowledge that an unintended audio-design issue likely shaped the immersion scores across all conditions, and that the marker manipulation was therefore competing against this baseline rather than operating on a neutral one. The second concerns future research and design practice: VR experiments testing narrative-immersion-related variables should consider audio mechanics as a potential confound to be controlled for or explicitly varied, and VR experiences intended for general or mixed-neurotype audiences should treat sustained or layered audio as a design parameter requiring deliberate calibration rather than as a neutral background. The specific finding that the xylophone's sustained playback degraded immersion for at least one participant who explicitly identified the mechanism suggests that even within the present study's sample, audio design choices were doing more interpretive work than the experimental design accounted for.

5.6 THE EXPLORATORY BEHAVIOUR OBSERVATION

Among the three themes identified in the qualitative analysis chapter, the third concerned a pattern in participant behaviour that emerged not from questionnaire responses or post-session conversations, but from the researcher's observations during testing. As reported in section 4.3.3, participants in the control and non-diegetic conditions appeared to attempt interaction with non-interactable objects - chairs, the bed, the door - more frequently than participants in the diegetic condition, who seemed to confine their interaction attempts more closely to the marked toys. This observation rests on informal impressions made during testing rather than on systematic behavioural recording, and no structured count of interaction attempts was maintained. It is examined in this chapter not as a finding the present study can defend on its own, but because the qualitative data, when read alongside the observational pattern, suggests two plausible mechanisms worth articulating for future research to test.

The first reading is consistent with the affordance framework established in the literature review. Norman's (2013) account holds that effective signifiers indicate not only what is interactable but, by implication, what is not - players who perceive a signifier marking a particular object should infer that unmarked objects do not afford the same interaction. From this perspective, the apparent reduction in static-object interaction attempts in the diegetic condition may reflect the brightness shader functioning as a signifier in both directions: highlighting the interactable toys and, by contrast, communicating non-interactability for the unmarked objects around them. The non-diegetic exclamation mark, despite being rated as the clearest signifier across the qualitative data, may not have produced the same effect because the importance of the marking system itself invited rather than discouraged probing of the unmarked items. A participant in the non-diegetic condition, working professionally as a sound designer, articulated exactly this dynamic: "it worked well to indicate what was interactable but I felt myself trying to pick up other things out of curiosity". This comment is striking because it confirms, from a participant's own perspective, that the explicit marking system did not suppress exploratory behaviour - and may even have invited it, by making the marked/unmarked distinction visually salient enough to become an object of curiosity in itself.

A second explanation attributes the same observational pattern to participant unfamiliarity with the medium rather than to signifier mechanics. Sample composition data reported in section 4.1.1 indicates that 48.6% of participants had no prior experience with VR games and an additional 34.3% reported low VR experience. For participants encountering VR for the first or near-first time, the medium itself presents an open question - "what is and is not possible to interact with here?" - that may motivate exploratory contact with the environment independent of any marker system. A participant in the non-diegetic condition, a literature student with no prior gaming experience, expressed this directly: "(...) it was not clear to me what I could grab and what I couldn't. For example I wanted to grab the chair but didn't know I could not. Maybe it's because I almost have no experience in gaming". This comment is informative on two levels. First, it indicates that the participant attributed their exploratory behaviour not to the marker design but to their own unfamiliarity with the medium - even in a condition with the most explicit marking system tested. Second, it suggests that VR-novice participants may bring a different orientation to the environment than experienced players. This pattern is consistent with the well-documented "novelty effect" in VR research, defined as the additional cognitive load imposed on users by new and unfamiliar ways of interacting with an immersive environment (Huang, 2003; Miguel-Alonso et al., 2024). Novice VR users, in particular, have been observed to focus on exploring the environment and testing all possible interactions rather than on the experience's intended content (Miguel-Alonso et al., 2024) - a description that maps closely onto the literature student's account of wanting to grab the chair to discover whether it was interactable.

These two readings - signifier-effectiveness versus medium-novelty curiosity - are not mutually exclusive, and the qualitative data is consistent with both operating simultaneously across different participants. What the observational pattern and the participant comments share is a recognition that interaction behaviour in VR is shaped by factors operating below the level of what validated immersion scales are designed to capture. Moment-to-moment exploratory behaviour, attentional allocation across marked and unmarked elements, and the orientation participants bring to an unfamiliar medium all influence how players engage with a VR environment, and none of these were directly measured by the NES or IEQ-SF. It is also worth noting that the post-session conversations did not surface frustration in any condition, and the participant who wanted to grab the digital chair did not describe the experience as immersion-breaking - they simply noted the ambiguity and offered an explanation for their own behaviour. If exploratory probing of static objects had degraded narrative immersion in a measurable way, the validated scales would have been positioned to detect it; that they did not is consistent with the reading that this kind of behaviour functions as orientation to the medium rather than as a signification failure.

A productive direction for future research follows from these observations. Behavioural logging within the game engine - recording the precise objects participants attempted to interact with, the timing of those attempts, and their distribution across the experience - would allow the informal pattern reported here to be tested systematically. Such measurement would also allow researchers to distinguish empirically between the two readings proposed above, since signifier-driven and curiosity-driven exploration would produce different temporal and spatial patterns of interaction attempts: signifier-driven probing would be expected to concentrate early in the experience as participants establish the marking system and then decline, while curiosity-driven probing would persist as the unfamiliar medium continues to invite testing. Stratifying analysis by prior VR experience would further refine this distinction, since the curiosity reading predicts that novice players would show different patterns of static-object

interaction than experienced players regardless of marker condition. This methodological extension is consistent with the broader argument advanced earlier in this chapter about the limits of adapted screen-based instruments in VR research, since behavioural measures operating directly on participant action would not require the cross-medium construct adaptation that the IEQ-SF struggles with.

5.7 PRACTICAL/INDUSTRY IMPLICATIONS

The introduction to this thesis stated that the research would aim to provide evidence-based guidelines for designing UI markers in narrative VR experiences. The null quantitative result complicates that intention, since it does not allow the present study to recommend one marker type over another on the basis of measured narrative immersion alone. What the study does allow, when its quantitative, qualitative, and interpretive findings are read together, is the articulation of a more nuanced set of design considerations than the binary "diegetic versus non-diegetic" framing of the existing literature tends to produce. The following implications are offered as considerations grounded in the present study's data and interpretations rather than as prescriptions established by it, and respect the limits imposed by the sample size, the single case study, and the moderating factors examined throughout this chapter.

The first implication concerns an assumption that runs through the literature reviewed in Chapter 2 and through design discourse more broadly: that non-diegetic UI elements risk disrupting narrative immersion by exposing what Murray (1997) called the "raw materials of creation", and that diegetic alternatives should therefore be preferred in narrative-driven VR experiences. This assumption, formalised in Hypothesis 2 of the present thesis and supported in prior empirical work by Iacovides et al. (2015) and Salomoni et al. (2016), did not bear out in the data collected here. The present study's findings do not support this assumption in the narrative VR context tested. Non-diegetic markers were not associated with lower narrative engagement scores than diegetic markers, were rated by participants as the clearest signifiers, and were the only marker type for which a measurable benefit emerged - the Markers_6 trend suggesting they may have aided participants in identifying which environmental elements carried narrative weight. For practitioners, this suggests that the choice between diegetic and non-diegetic markers in narrative VR can be made on grounds of clarity, recognisability, and salience signalling rather than on the assumption that immersion will automatically be lost with the introduction of conventional signs. The exclamation mark used in the present study was stylistically integrated into the child's-bedroom aesthetic - drawn in the same crayon-like style as the rest of the environment - and several participants explicitly commented on this stylistic integration. Designers working with non-diegetic markers in narrative VR may therefore find more benefit in attending to the aesthetic coherence of those markers with the surrounding world than in avoiding non-diegetic markers altogether.

A second implication concerns the role of natural object affordances. As discussed in section 5.3, the control condition's narrative immersion scores were closely comparable to those of the two marker conditions, and at least one participant explicitly recognised the unmarked interactable toys as such on the basis of their visual properties alone. For designers, this suggests that the design of the interactable objects themselves carries significant signification work that designed markers must compete with or amplify. Where interactable objects are visually distinct from non-interactable surroundings - through scale, material, animation, or stylistic difference - additional marking systems may add less value than designers commonly assume. Conversely,

where interactable objects are visually similar to non-interactable surroundings, designed markers are likely doing more of the signification work, and their design choices become correspondingly more consequential. This recommendation extends Steffen et al.'s (2019) account of inherited VR affordances into a practical heuristic: before adding a marker system, designers might productively ask whether the objects themselves already communicate what the marker is intended to communicate.

A third implication concerns the relationship between UI design and the sensory channel through which a VR experience delivers its narrative. The interpretation advanced in section 5.1.1 - that marker type may matter most when the UI and the narrative content compete for the same sensory channel, and less when they operate on separate channels - has direct application to design decisions. For visually-driven narratives, where the story is told through environmental events, text, or character animation, the visual coherence of the marker system with the narrative content is likely to influence immersion more substantially than in audio-driven narratives. For audio-driven narratives, where the visual marker system and the narrative content occupy different channels, designers may have more latitude in marker choice without measurable cost to narrative engagement. This consideration also extends to the audio design of the experience itself. The sound design finding discussed in section 5.5 suggests that mechanical sounds produced by interaction (in the present study, the musical sounds of the toys) can compete with audio narrative for the same attentional bandwidth. Designers working on narrative VR experiences with audio-led storytelling should treat the interaction-mechanical audio not as a neutral feedback layer but as a design parameter that requires deliberate calibration against the narrative audio it shares the channel with.

A fourth implication concerns the VR-novice user. As discussed in section 5.6, participants with limited prior VR experience appeared to engage in exploratory probing of static objects independent of marker condition, a pattern consistent with the documented "novelty effect" in VR research (Miguel-Alonso et al., 2024). For the rapidly growing population of first-time VR users that the commercial VR market increasingly addresses (Kemp, 2025), this has two implications for designers. First, marker systems should not be designed on the assumption that the absence of a marker on an object will reliably communicate non-interactability to novice users, whose orientation to the medium itself may motivate exploratory contact regardless of signification. Designers may therefore wish to consider how non-interactable objects respond - or fail to respond - to attempted interaction, since this response constitutes a signification moment in its own right, particularly for novice players. Secondly, tutorial design and onboarding for VR-novice users remain important. Empirical work has shown that structured VR tutorials can meaningfully mitigate the novelty effect by helping first-time users gain familiarity with controllers and interaction conventions before engaging with the experience's main content (Miguel-Alonso et al., 2023). For narrative VR experiences specifically, this consideration is doubly important: a player still orienting to what the medium affords has less attentional capacity available for the story itself, so reducing that early-experience cognitive load may free attentional resources for engagement with the narrative content.

A fifth and final implication concerns accessibility considerations that emerged from the discussion of sound design in section 5.5. The auditory hyper-reactivity literature reviewed in that section suggests that layered or repetitive sound design may impose disproportionate cognitive cost on neurodivergent users, who are present in any general-population sample even when not explicitly recruited for. For commercial VR experiences intended for broad audiences, this points toward providing user-adjustable audio parameters - allowing players to mute,

reduce, or simplify the interaction-mechanical audio independent of the narrative audio - as a relatively low-cost design feature that meaningfully expands the population of users who can engage with the experience without sensory overload. This recommendation is consistent with the broader direction of inclusive design practice in VR (Hulusic et al., 2021) and aligns with the increasing emphasis on accessibility as a baseline expectation rather than an optional feature in contemporary game development.

Taken together, these implications reframe what the present study contributes to design practice. Rather than recommending one marker type over another, the study contributes a set of considerations that direct the designer's attention toward factors that the diegetic/non-diegetic framework alone obscures: the natural affordances of the objects being marked, the sensory-channel relationship between UI and narrative, the orientation that VR-novice users bring to the experience, and the accessibility implications of audio design choices. For practitioners working on narrative VR experiences, these considerations may prove more actionable than a binary recommendation, since they direct design thinking toward the specific conditions under which different marker strategies are likely to succeed or fail.

5.8 ENGAGEMENT WITH THE ORIGINAL RESEARCH GAP

The literature review of this thesis identified a specific gap in the existing VR UI research: prior empirical studies of diegetic versus non-diegetic UI in VR - notably Saling et al. (2021) and Köhle et al. (2021) - had focused exclusively on shooter contexts and measured spatial presence or user preference, leaving the question of how UI design influences narrative immersion in story-driven VR empirically untested. The literature review further hypothesised that narrative immersion might respond differently to UI design than spatial presence does, on the grounds that narrative immersion depends on sustained cognitive and emotional engagement with story content rather than on the hardware-mediated sense of "being there" that VR headsets reliably produce. The present study was designed to address this gap directly, by transferring the experimental comparison from action-oriented gameplay to a narrative-driven VR experience and measuring narrative immersion specifically rather than presence. The following discussion considers what the field's understanding looks like now that this study has contributed its findings.

The most direct contribution is methodological. To the researcher's knowledge, the present study is the first to test the diegetic/non-diegetic UI comparison in a narrative-driven VR experience using validated narrative engagement and immersion instruments, with a third control condition included to establish a baseline. This design choice extends the comparative work of Saling et al. and Köhle et al. into a genre where their methodology had not previously been applied, and answers the literature review's empirical question - what happens when this experimental comparison is run in a narrative VR context - in concrete terms. The answer the data provides is that, in the conditions tested, marker type does not produce measurable differences in narrative immersion as operationalised by the NES and IEQ-SF. This finding addresses the gap as it was originally framed.

The substantive contribution is more nuanced and, in many ways, more theoretically interesting. The hypothesis underlying this thesis was that narrative immersion would behave differently from spatial presence in response to UI manipulation - more sensitive, more vulnerable to disruption, more responsive to the diegetic/non-diegetic distinction. The data did

not bear out this hypothesised distinction. Across three independent VR UI studies - Saling et al. on spatial presence, Köhle et al. on user preference, and the present study on narrative immersion - no significant effect of UI type has been observed. This convergent pattern of null findings, read across three constructs that the literature treats as distinct, suggests that the diegetic/non-diegetic framework as inherited from Fagerholt and Lorentzon's (2009) screen-based gaming context may not transfer to VR in the way the framework's continued use in the field implies. The contribution of the present study to the gap, then, is not only the answer "no measurable effect on narrative immersion" but also the broader, more cautious claim that the framework around which VR UI research has been organised may need rethinking. This is a contribution that none of the three studies could have made alone but that the cumulative pattern across them now makes available.

Beyond closing the gap as originally framed, the present study has identified several adjacent gaps that the research process itself surfaced. The first concerns measurement: section 5.2 of this chapter demonstrated that the IEQ-SF's reliability dropped substantially when transferred from screen-based gaming to VR, particularly on the Real-World Dissociation subscale. This finding contributes to a methodological gap that the field has acknowledged but not yet acted on - the absence of VR-native immersion instruments validated for use within the medium rather than adapted from it. The second concerns the sensory-channel relationship between UI and narrative content: section 5.x advanced the hypothesis that marker type may matter most when the UI and the narrative content compete for the same sensory channel and matter less when they operate on separate channels. This hypothesis, if borne out by further research, would refine the diegetic/non-diegetic framework rather than replace it - identifying the conditions under which the framework's predictions are most likely to apply. The third concerns the role of natural object affordances, examined in section 5.x: the present study's data suggests that the affordances objects carry by virtue of their visual and physical properties may do more signification work than the existing UI research has accounted for, and that future experimental designs need to treat natural affordance as a distinct condition rather than as a no-signifier baseline.

A fourth direction emerged from the qualitative and observational data on VR-novice participants. The present study's sample included a high proportion of first-time and low-experience VR users (over 82% combined), and both the questionnaire responses and the observational notes pointed toward exploratory behaviour shaped by the novelty effect documented in the broader VR literature (Miguel-Alonso et al., 2024). The interaction between VR novelty and UI signification remains underexplored: it is plausible that signifier-effectiveness research conducted predominantly with VR-experienced participants produces different findings than the same research conducted with general-population samples, and the field has not yet systematically tested this. As VR continues to expand into broader consumer markets (Kemp, 2025), the question of how marker design interacts with VR familiarity becomes increasingly consequential for both researchers and designers.

Taken together, the present study contributes to the literature on three levels. It closes the gap it set out to close, by providing empirical evidence on diegetic versus non-diegetic UI in a narrative VR context. It contributes a broader theoretical observation - that the framework's continued utility in VR is open to question - that the cumulative pattern of null findings across the field now supports. And it identifies several adjacent gaps - measurement, sensory channel mediation, natural affordances, and VR-novice user behaviour - that future research can

productively address. The conclusion chapter will return to several of these directions as concrete future-research recommendations.

5.9 GENERALISABILITY

The findings of this study apply to a specific set of conditions, and articulating those conditions clearly is essential to readers and future researchers seeking to build on the work. This section discusses the scope within which the present findings can be reasonably extended, organised around three dimensions: the population studied, the stimulus material used, and the constructs through which immersion was operationalised. The intent is not to repeat the limitations already acknowledged in the methodology chapter (section 3.6), but to consider what those limitations imply for the inferential reach of the study's conclusions.

The first dimension concerns the population. Findings extend to adults aged 19-39 residing in the Netherlands, able-bodied, English-speaking, and with the physical mobility required for a standing VR experience. As acknowledged in the methodology chapter, this target group does not represent the general population. The sample of 35 participants was further shaped by the convenience and purposive sampling strategy used to recruit it, which produced a sample weighted toward game-related (34.3%) and creative (22.9%) professional backgrounds. Combined, these two categories accounted for 57.2% of participants - meaning that more than half of the study's participants came from fields in which visual literacy and exposure to game design conventions are professionally cultivated. Findings should not be assumed to extend to general-population samples without targeted replication, since a participant population with greater familiarity with design conventions and visual signification may respond differently to marker design choices than a less design-literate population would. The same caveat applies to demographic extension: results may not generalise to children, older adults, or populations outside the Netherlands and the cultural context in which the study was conducted. The reverse caveat is also worth stating: the high proportion of VR-novice participants (over 82% combined no-experience and low-experience users) may make some findings *more* rather than less generalisable to first-time VR users, but limits their extension to experienced VR populations whose orientation to the medium differs systematically.

The second dimension concerns the stimulus material. The findings of this study are grounded in a single narrative VR experience - "Echoes of Abandonment" - with specific characteristics that bound their applicability. The experience delivers its narrative primarily through audio (parents arguing behind a door), uses a clay-sculpted environmental aesthetic rendered through photogrammetry, lasts approximately seven to ten minutes, and centres on a single interaction type: pick-up of toy objects within a child's bedroom setting. The two designed marker conditions are equally specific - a brightness shader applied to interactable objects and a pulsating exclamation mark icon styled to resemble a crayon drawing. The null finding on marker type, the qualitative pattern around non-diegetic marker clarity, the sound design observation, and the natural affordances observation are all conditional on these stimulus characteristics. Findings should not be assumed to extend to action-oriented VR genres (shooters, sports, simulation), to social or multi-user VR contexts, to longer-form VR experiences in which novelty effects may dissipate over time, or to VR interactions beyond pick-up (locomotion, manipulation, gesture-based input). The sensory-channel hypothesis advanced in section 5.x - that marker type may matter most when UI and narrative compete for the same sensory channel - specifically does not extend automatically to visually-driven VR narratives;

testing that hypothesis there is identified as a future research direction rather than a finding the present study has established.

The third dimension concerns the constructs through which narrative immersion was measured. As discussed at length in section 5.2, the present study operationalised narrative immersion through two validated instruments: the Narrative Engagement Scale and the Immersive Experience Questionnaire - Short Form. The NES held up well as a measurement instrument in this study, with reliability comparable to its performance in screen-based contexts; the IEQ-SF performed substantially less well, with Real-World Dissociation reaching only $\alpha = .439$. Findings should be understood as applying to immersion as operationalised by these scales, not as applying to immersion-as-such. Other constructs that the present study did not measure - moment-to-moment perceptual clarity, attentional allocation across marked and unmarked elements, embodied presence, social presence, flow - might respond to marker design in ways the NES and IEQ-SF were not positioned to detect. The qualitative findings reported in section 4.3 strongly suggest that meaningful experiential differences existed between conditions even where the validated scales did not record them; future research using complementary or VR-native measurement approaches may therefore find effects that the present instrumentation did not surface.

Beyond these three dimensions, two broader scope conditions apply. The study was conducted in late 2025 on Oculus Meta 3 hardware, and findings are bounded by the hardware capabilities and user population available at that moment; as VR hardware evolves and as the consumer VR population shifts in composition (Kemp, 2025), the findings' applicability to future research and practice may shift accordingly. Additionally, the study examined diegetic and non-diegetic markers only, leaving the other two categories of Fagerholt and Lorentzon's (2009) framework - meta and spatial UI - outside the scope of comparison. Findings about diegetic versus non-diegetic markers should not be extrapolated to claims about UI design in VR more broadly.

Taken together, the present study's findings are most defensibly applied to: young Dutch adults with limited VR experience, playing short-form audio-driven narrative VR experiences with pick-up interactions, where narrative immersion is measured through screen-based-validated scales. Within these bounds, the findings provide the empirical evidence and theoretical interpretations developed throughout this discussion chapter. Beyond these bounds, they should be treated as hypotheses to be tested by future research rather than as conclusions to be assumed.

5.10 Ethical implications

The methodology chapter (section 3.8) and the results chapter (section 4.6) addressed the ethical considerations that shaped the design and conduct of this study: physical safety and motion sickness risk, mental wellbeing in relation to the narrative content, social desirability effects arising from the researcher's presence during testing, GDPR-compliant data handling, and the representational implications of recruiting from design-literate academic networks. These considerations were anticipated before data collection and addressed through procedural choices made at the methodology stage. The present section examines three additional ethical considerations that the research process and the interpretive work of this discussion chapter brought into view, and which the methodology chapter could not have addressed because they only became visible later: the implications of the study's audio design for neurodivergent

participants, the trade-offs embedded in the choice not to audio-record qualitative data, and the forward-looking ethical implications of how the study's findings may be used in design practice.

5.10.1 Neurodivergence

The first consideration arose from the interpretive work of section 5.1.4. As discussed there, the layered sound design of "Echoes of Abandonment" - particularly the sustained xylophone playback and the simultaneous toy sounds competing with the parents' argument - may have imposed disproportionate cognitive cost on participants with auditory hyper-reactivity, a documented feature of both autism spectrum and ADHD presentations. Dutch population data indicates that approximately 3% of adults self-report autism spectrum disorder (CBS, 2025) and 1-2.5% meet diagnostic criteria for ADHD (de Graaf et al., 2014); both figures may be lower than actual prevalence, since many adults remain undiagnosed. The consent form used in this study informed participants that the experience involved standing VR use of approximately ten minutes and that the audio content might be emotionally uncomfortable, but it did not specifically flag the layered audio mechanic as a potential source of sensory overload for neurodivergent participants. In retrospect, this represents an ethical gap: participants for whom the audio design produced sensory discomfort had no advance warning specific to that aspect of the experience, and may have continued the experience past their comfort threshold in part because the comfort issue was not one the consent form had prepared them to anticipate. The ethical responsibility of the present study cannot be undone retrospectively, but the consideration has direct implications for future research and design practice. Studies involving layered, repetitive, or sustained audio mechanics should treat sensory accessibility as a distinct ethical consideration rather than as a sub-case of general physical wellbeing, and consent procedures should specifically describe the audio characteristics that may affect neurodivergent users so that participants can make informed decisions about their participation.

5.10.2 Undocumented qualitative data

The second consideration concerns the deliberate choice not to audio-record post-session conversations. As noted in section 4.1, qualitative data from these conversations was captured through the researcher's notes taken shortly after each session rather than through verbatim recording. This methodological choice carried both ethical and analytical consequences worth examining. On the ethical side, the absence of audio recording reduced the perceived stakes of participation: participants spoke freely in casual post-session conversations without the implicit pressure that microphones impose, lowered consent-burden by avoiding additional permissions for audio storage and retention, and protected participants from the small but real risk of voice-identifiable data being mishandled or accessed beyond the research context. These are real ethical benefits, not incidental ones. On the analytical side, however, the same choice means that the qualitative findings reported in this thesis represent the researcher's interpreted reconstruction of participant feedback rather than verbatim records, and should be weighted accordingly - an acknowledgment already made in section 4.1. The ethical trade-off worth naming explicitly is that prioritising participant comfort and data minimisation came at the cost of evidential precision, and that this is a defensible trade-off rather than a methodological oversight. Future research with greater resources for participant support might productively explore middle-ground approaches: brief structured interviews with audio recording and explicit consent, voluntary participation in a recorded follow-up, or anonymised transcription that decouples voice from response. The present study's approach was appropriate to its scale

and scope; future studies seeking stronger qualitative evidence should weigh the additional consent burden against the analytical benefit and document that decision transparently.

5.10.3 Open ended use of findings

The third consideration is forward-looking and concerns the potential applications of the present study's findings beyond the context in which they were generated. The introduction to this thesis framed the research as contributing to the design of narrative VR experiences, and the practical implications laid out in section 5.7 are directed primarily toward designers working on story-driven VR for entertainment and creative-cultural contexts. However, the underlying findings - that non-diegetic markers do not necessarily disrupt narrative immersion, that natural object affordances do significant signification work, that audio-led narrative can coexist with visual UI without measurable cost to story engagement - are not restricted in their applicability to entertainment contexts. The same findings could inform the design of VR experiences intended for advertising, persuasive communication, attitudinal change, or behavioural conditioning, where the maintenance of narrative immersion serves goals other than artistic engagement. Narrative transportation, the construct from which this thesis's measurement approach is derived (Green & Brock, 2000), has been empirically associated with reduced critical evaluation of message content - meaning that the same design choices that preserve immersion in an artistic narrative may also reduce the critical distance available to a player encountering a persuasive narrative. This is not a hypothetical concern: VR is increasingly being adopted in marketing, corporate training, healthcare communication, and political messaging contexts, where the design choices examined in this study could be applied. The present study does not address the ethics of those applications and cannot prescribe how its findings should or should not be used by practitioners. What it can do, in the spirit of Linhares Pontes and others who have argued for designer accountability in immersive media, is name the consideration: the affordance-communication and immersion-preserving strategies investigated here are tools, and tools developed for one purpose can be applied to others. Researchers and designers working with narrative VR techniques should remain attentive to the contexts in which immersion-supporting design choices are deployed, and to whether those contexts respect the player's capacity to engage critically with the content being presented.

Taken together, these three considerations extend rather than replace the ethical groundwork laid in the methodology and results chapters. The first acknowledges an ethical gap that became visible only through the interpretive work of this discussion: that the study's audio design carried implications for neurodivergent participants that the consent process did not specifically address. The second names a methodological trade-off as the ethical decision it actually was: prioritising participant comfort and data minimisation at the cost of evidential precision. The third looks forward, recognising that the present study's contributions enter a research landscape and design practice in which they may be applied beyond the contexts the researcher anticipated. None of these considerations invalidate the study's findings, but each shapes how the findings should be carried forward by the researcher, by future researchers building on this work, and by practitioners drawing on its design recommendations.

6 CONCLUSION

This thesis investigated the research question *"What is the influence of interaction markers design in VR games on the player's ability to become immersed in narrative content among young adults living in the Netherlands?"*. The motivation for this question was laid out in the introduction: as VR matures into a mainstream platform for interactive storytelling (Irshad & Perkis, 2020), the design of its symbolic interfaces, and in particular of the markers that signal which 3D objects can be interacted with, becomes consequential for whether the medium's narrative potential is realised or undermined. Prior empirical work on diegetic versus non-diegetic UI in VR (Saling et al., 2021; Köhle et al., 2021) had examined this question in shooter contexts and measured spatial presence or user preference, leaving the relationship between marker design and *narrative* immersion in story-driven VR experiences empirically untested. The present study addressed that gap by testing three marker conditions: diegetic, non-diegetic, and a no-marker control; within a single narrative VR experience, *Echoes of Abandonment*, with 35 participants and two validated measurement instruments supplemented by custom Likert items.

The quantitative analysis returned null findings across all measured constructs. The one-way ANOVA on NES total scores produced no statistically significant difference between conditions, and the same pattern held across the NES subscales, the IEQ-SF, and its subscales (chapter 4.2). Effect sizes were negligible throughout (all $\eta^2 < .02$). The qualitative analysis, summarised in chapter 4.3, identified meaningful perceptual differences that the validated scales did not capture: non-diegetic markers were described as the clearest signifiers across conditions, the layered audio design of the experience was reported as cognitively demanding by several participants, and at least one control-condition participant identified the unmarked toys as interactable on the basis of their visual properties alone. The custom item Markers_6 ("The interaction markers helped me understand what was important to the story") produced the closest result to statistical significance in the entire dataset ($p = .087$), with the non-diegetic group scoring notably higher than the other two, what is a pattern that runs orthogonally to the hypotheses this thesis set out to test.

In direct answer to the research question, the present study finds no measurable influence of marker design on narrative immersion under the conditions tested. Hypothesis 1: that diegetic markers would support narration immersion, was not supported, since the diegetic condition did not produce higher NES scores than the control condition. Hypothesis 2: that non-diegetic markers would noticeably disturb narration immersion, was not supported either; if anything, the non-diegetic condition emerged from the qualitative and custom-item data as the clearest signifier, with no detectable cost to story engagement. This answer should be read as conditional rather than absolute, since the study's sample size, single case study design, and audio-primary narrative delivery each impose interpretive bounds that chapter 5 examined in detail.

The contribution this thesis makes to the field operates on two levels. The direct contribution is empirical: to the researcher's knowledge, the present study is the first to test the diegetic/non-diegetic UI comparison in a narrative-driven VR experience using validated narrative engagement and immersion instruments, with a no-marker control condition included as a

baseline. This extends the comparative work of Saling et al. (2021) and Köhle et al. (2021) into a genre where their methodology had not previously been applied. The broader contribution is theoretical, and it emerges from the cumulative pattern of null findings now visible across three independent VR UI studies - two on spatial presence and user preference, one on narrative immersion that have all failed to detect significant effects of UI type on their measured outcomes. Taken together, these convergent null results suggest that the diegetic/non-diegetic framework, as theorised by Fagerholt and Lorentzon (2009) for screen-based games, may not transfer to VR in the way the framework's continued use in the field implies, a claim no individual study could make alone but that the cumulative pattern now supports. Beyond closing the original research gap, the present study also surfaced several adjacent gaps the field has not yet acted on: the reliability concerns observed with the IEQ-SF in VR contexts (chapter 5.2), the role of natural object affordances as a third form of signification rather than a no-signifier baseline (chapter 5.3), and the question of how the sensory-channel relationship between UI and narrative content mediates marker effectiveness (chapter 5.1.2).

Several limitations bound these contributions and have been discussed in detail across chapters 3.6, 4.5, and 5.10. The sample of 35 participants, while consistent with comparable VR studies, limits the statistical power available to detect small effects and was further shaped by a convenience-sampling strategy that produced a participant group weighted toward game-related and creative professional backgrounds. The findings are grounded in a single narrative VR experience with specific audio-primary, pick-up-based, short-form characteristics, and should not be assumed to extend to other VR genres or interaction types without targeted replication. The IEQ-SF showed substantially weaker reliability in this VR context ($\alpha = .487$) than the NES did ($\alpha = .869$), which means the IEQ-SF subscale results are best read as consistent with a genuine null effect rather than as strong evidence of one. Two further considerations became visible only through the interpretive work of chapter 5: the study's layered audio design carried implications for neurodivergent participants that the consent process did not specifically address, and the deliberate choice not to audio-record post-session conversations means the qualitative findings represent the researcher's reconstruction of participant feedback rather than verbatim records, a defensible trade-off rather than an oversight, but one that future research with greater resources may want to revisit.

From these limitations and from the adjacent gaps the present study surfaced, several directions for future research follow, and are developed in detail in chapter 7. Most pressingly, the field would benefit from the development of VR-native immersion instruments rather than continued reliance on scales adapted from screen-based media (chapter 7.1). Beyond measurement, future research could productively test whether the sensory-channel hypothesis advanced in chapter 5.1.2 holds in visually-driven rather than audio-driven VR narratives, treat natural object affordance as a distinct experimental condition rather than as a no-signifier baseline (chapter 5.3), and examine how marker design interacts with VR familiarity as the consumer VR population continues to expand into general audiences (Kemp, 2025). For practitioners working on narrative VR experiences, the present study offers not a binary recommendation but a set of considerations developed in chapter 5.7: that non-diegetic markers can be used in narrative VR without the immersion cost the literature has tended to assume, provided their aesthetic is integrated with the world they inhabit; that the natural affordances of interactable objects do significant signification work that designed markers must

compete with or amplify; and that the audio design of an interaction-led VR experience requires deliberate calibration against the narrative audio it shares a channel with.

The central point this thesis asks its readers to take away is therefore not that marker design does not matter for narrative immersion in VR, but that the framework through which the field has been asking the question may be too coarse to detect what matters within it. The diegetic/non-diegetic distinction inherited from screen-based gaming foregrounds one dimension of UI design, its position relative to the diegesis, while leaving other dimensions that the present study's data identifies as consequential largely unexamined. As VR continues to grow as a platform for interactive storytelling, the field's understanding of how its interfaces support or disrupt narrative engagement will be strengthened more by refining the questions it asks than by accumulating null results to questions whose framing the medium has outgrown.

7 FUTURE DIRECTIONS

7.1 DEVELOPMENT OF BETTER SCALES FOR VR RESEARCH

As VR matures as both a consumer technology and a research medium, the field is reaching a point where adapted instruments are insufficient and purpose-built ones become necessary. Several concrete directions emerge for the development of VR-native immersion measurement. First, future scales should reconceptualise rather than simply re-validate the Real-World Dissociation construct, since the dimension that screen-based scales treat as a continuous variable is, in VR, a near-constant condition imposed by the hardware. A VR-native equivalent might instead measure the *strength* of presence within the virtual environment, the *moments* at which awareness of the physical world breaks through, or the participant's *capacity to be drawn out* of the virtual space — distinct constructs that the current Real-World Dissociation items conflate. Second, scales should distinguish more carefully between spatial presence and narrative immersion as separate dependent variables; the existing literature, including Saling et al. (2021) and Köhle et al. (2021), tends to use these terms with significant overlap, which obscures cases like the present one where the two constructs may respond differently to the same manipulation. Third, self-report instruments could be productively complemented by behavioural and physiological measures less vulnerable to scale-design assumptions: interaction logging within the game engine (already recommended as a methodological improvement in section 4.3.3), eye-tracking data from headsets that increasingly include the capability natively, and post-experience structured interviews designed to capture moment-to-moment perceptual experiences that quantitative scales appear to miss. Taken together, these directions describe not a single replacement instrument but a *measurement toolkit* tailored to the specific affordances and constraints of the VR medium.

7.2 AUDIO IN VR RESEARCH

The finding also enters into conversation with the theoretical model proposed by Pillai and Verma (2019), discussed in section 2.3.2 of the literature review. Their account of narrative immersion in VR cinema distinguishes between technical immersion (sustained through attentional cues and

environmental fidelity) and narrative immersion (arising from story, character, and viewer integration), and argues that these two layers can support and reinforce each other to strengthen the overall immersive experience. The present study's qualitative data does not contradict that claim, but it does suggest a condition under which their model may not hold: when the technical or mechanical layer of a VR experience produces sound that competes for the same sensory channel as the narrative content. In "Echoes of Abandonment", the sound-producing toy mechanic and the audio-delivered narrative occupied the same auditory bandwidth, and rather than reinforcing each other in the manner Pillai and Verma describe, they appear to have created interference. This suggests that the supportive relationship between technical and narrative immersion that their model proposes may be conditional on the two layers operating either on different sensory channels or on the same channel without competing for attentional focus. This is a refinement of their model rather than a refutation, but it is the kind of refinement that the field of narrative VR research has not yet articulated clearly.

7.3 DIEGETIC/NON-DIEGETIC DIVISION MIGHT NOT TRANSFER TO VR

APPENDIX A

SCRIPT PARENTS' FIGHT IN THE GAME "ECHOES OF ABANDONMENT"

DAD

Honey, I'm home.

MOM

How was work?

DAD

Ah, exhausting. I'm famished, what's for dinner?

MOM

We already ate dinner, but we have some frozen fish sticks you can warm up for yourself.

Dad sighs.

MOM

What?

DAD

Nothing.

MOM

No, c'mon. What is it?

DAD

I've been at work all day. I come home to a mess and apparently I can't even expect warm food on the table. That's the least I could hope for when I'm breaking my back for this family.

MOM

YOU'RE the one breaking your back? I'm the one helping our kid with their homework, getting them ready for school, and feeding them. I'm washing everyone's clothes. I'm basically raising our child by myself. So sorry, if there's a single sock on the carpet that your OCD can't handle.

DAD

You don't have any other tasks now than this child, so I don't see what the hassle is. It's so little, yet all you can do is fuck it up.

MOM

Don't believe anybody could beat you at fucking things up. Literally one thing you needed to do for the kid was to take them to the dentist last week and of course you forgot!

DAD

I thought it was a day later cut me some slack! I'm working overtime and going on business trips for both of you !

MOM

Do you think I'm stupid? Don't you think I know what your 'business trips' are? Leaving for damn vacations with your loser friends when you have a family at home. You don't have the right to tell me shit. You're just looking for any reason not to come home. And our kid needs both parents.

DAD

Don't use them against me! There isn't much to come home to evidently. Why would I want to spend more time than I have to in this miserable fucking house? Eating your miserable frozen food? And seeing your miserable fucking face?

MOM

Maybe I wouldn't be so miserable if you were more of a man and less of a self-absorbed cunt!

The child walks tentatively towards the door to hear the talk more clearly, but accidentally knocks over a tower of toy building blocks. The argument stops momentarily. The parents start to whisper.

DAD

(whispers aggressively through his teeth)

See what you did?! You woke the child up.

MOM

Maybe we should tell them about the divorce.

DAD

Oh just like that?

MOM

Yeah, just like that, it's mostly you fault not to put any effort in our relationship so at leat you can do is to tell them the truth, I'm done pretending everything is fine.

DAD

Why do you always blame everything on me?? It's not my fault we're arguing all the time! And you want to take everything from me with this divorce too!

MOM

Oh, here we go again with the money talk. It's not about taking everything. It's about dividing what's rightfully ours.

DAD

Rightfully ours? You never contributed anything substantial to this marriage except spending my hard-earned money!

MOM

Spending it for OUR child you jerk! It's not like you ever helped me when I was going through a hard time after I gave birth. No emotional support whatsoever, just straight to work and I'm supposed to figure stuff out on my own!

DAD

Emotional support? Oh, please! Don't act like you're some saint in this marriage. You were always too busy with your friends or your hobbies to care about my needs! Someone needed to sustain our family financially, I had to work overtime. How do you think I'm feeling?!

MOM

Friends and hobbies? Maybe if you bothered to show any interest in me, I wouldn't have to seek companionship elsewhere!

DAD

[Shouting] Is that what this is about? Are you cheating on me?

MOM

[Tears welling up in her eyes] How dare you accuse me of that! You're the one who's never around, who's always finding excuses to avoid being with your family! No! I'm not cheating on you! It's about feeling neglected and unloved in this marriage!

DAD

Here you go again with manipulating me with those tears. Oh, so now it's my fault? Maybe if you were more satisfying to be around, I wouldn't need to escape! Another reason we should go our separate ways.

MOM

You call it escape? You're running away from your responsibilities!

DAD

And what about your responsibilities? All you do is nag and complain!

MOM

Then leave! See if I care! I can't do this anymore!

DAD

You wouldn't survive without me around and you know it.

Static noise starts*

APPENDIX B

FULL QUESTIONNAIRE USED IN THIS STUDY:

IEQ-SF:

1. I felt focused on the game

2. The game was something that I was experiencing, rather than just doing

3. I felt motivated when playing the game
4. I enjoyed playing the game
5. I felt consciously aware of being in the real world whilst playing
6. I forgot about my everyday concerns
7. I felt that I was separated from the real world environment
8. I found myself so involved that I was unaware I was using controls
9. I found the game challenging
10. I found the game easy
11. I found in suspense about whether or not I would do well in the game

NES:

12. At points, I had a hard time making sense of what was going on in the program.
13. My understanding of the characters is unclear.
14. I had a hard time recognizing the thread of the story.
15. I found my mind wandering while the program was on.
16. While the program was on I found myself thinking about other things. ()

17. I had a hard time keeping my mind on the program. ()

18. During the program, my body was in the room, but my mind was inside the world created by the story.

19. The program created a new world, and then that world suddenly disappeared when the program ended.

20. At times during the program, the story world was closer to me than the real world.

21. The story affected me emotionally.

22. During the program, when a main character succeeded, I felt happy, and when they suffered in some way, I felt sad.

23. I felt sorry for some of the characters in the program.

Custom questions:

24. What condition were you part of?

25. It was intuitive to understand how I could interact with objects in the virtual environment

26. Sounds that played when I interacted with the toys enhanced my enjoyment of the game

27. The marker of interactable toys made it easy for me to recognize which toys I can pick up

28. I found it easy to follow the story while playing the game

29. The interaction markers helped me understand what was important to the story.

30. What is your level of experience with VR games?

31. What is your profession / specialty?

32. What did you like about the design of the markers present in your VR experience?

APPENDIX C

CONSENT FORM OF THIS STUDY:

CONSENT FORM Master Thesis

RESEARCH TITLE: *Marking tangible user interface in VR games to support narrative immersion.*

I _____(your full name) provide consent to the following research instruction:

Breda University of Applied Sciences (BUas)

to use my data within the above-menoned project.

- I have been informed about the research project entitled '*Marking tangible user interface in VR games to support narrative immersion*' and have been requested to give consent in this form for the data collection and use of the 15-30 min long study including questionnaire, on _____ (today's date).
- I have been informed about the goal of this project and I have had the opportunity to ask the research team any questions which may arise about the research and my participation.
- I understand that my participation is voluntary and I am free to withdraw at any point during the play session and questionnaire without consequences. Because responses are collected anonymously without identifying information, withdrawal after questionnaire submission is not possible as your data cannot be identified. Data will be retained until the thesis is completed and graded, after which it will be securely deleted within 30 days.
- My refusal to participate or withdrawal of consent will not affect my position and my relationship with the institutions / researcher involved.
- I understand that the VR experience contains audio of parents arguing, which some participants may find emotionally uncomfortable. I am free to remove the headset and discontinue participation at any time without explanation or consequences.
- I understand this study involves standing while wearing a VR headset for approximately 10 minutes. I confirm I have no conditions (photosensitive epilepsy, severe motion sickness, mobility limitations) that would make this unsafe for me.

If I have any enquiries about the research, I can contact Paulina Cywoniuk (paulinacywoniuk@gmail.com), the responsible researcher.

By signing below, I am indicating my consent to:
• Provide access to the collected data regarding the questionnaire answers
• Having the collected data shared and to be used for research purposes and related publications.
• Being contacted if further clarification is necessary to investigate statements present in the collected data.
• All data presented in the thesis will be presented in an anonymized way.

I understand that the data collected from my participation will be used for the purpose of this research and will be securely stored, and I consent for it to be used in that manner.

Date: ..... Place: .....
Signature:
.....

APPENDIX D

EXPERIMENTS OF THE MARKERS DESIGN:

https://docs.google.com/document/d/1kEwaHedR5HL2x_Q3f4gRH2QuqczBap0VPFQsEFpRy4o/edit?usp=sharing

APPENDIX E

7.4 SCRIPT FOR THE RESEARCH PROCEDURE.

The following script describes the standardised verbal procedure used by the researcher with each participant. While the exact wording varied slightly between sessions, the content and sequence were consistent across all 35 participants.

7.5 PHASE 1: PRE-GAME SETUP

Upon entering the testing room, the participant was asked:

"Have you played a VR game before?"

Regardless of their answer, the researcher proceeded with headset installation. The participant was asked to hold the headset to their eyes in a comfortable position while the researcher secured the straps. Once the headset was fitted, the researcher provided movement instructions and the following verbal briefing:

"You are about to experience a piece of narrative that is delivered to you through audio, so please pay attention to what you are hearing."

The participant was then left to play the game without interruption, with the researcher observing quietly from a distance.

7.6 PHASE 2: POST-GAME CHECK

After the participant completed the VR experience (approximately 7–10 minutes), the researcher removed the headset and asked:

"Are you feeling dizzy, or experiencing any headache or symptoms of nausea?"

Participants were given time to sit down and rest before proceeding to the questionnaire.

7.7 PHASE 3: QUESTIONNAIRE BRIEFING

Before beginning the questionnaire, the researcher provided the following contextual briefing:

"You have just experienced a virtual reality narrative piece about parents arguing behind a door. You were playing the role of a child listening to this argument. By playing with the toys in your room, you were creating a musical composition that muted the argument of the parents, allowing you to symbolically reclaim your safe space. The purpose of the game was to interact with the toys."

The participant was then presented with the questionnaire on the Qualtrics platform.

BIBLIOGRAPHY

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly: Management Information Systems*, *24*(4), 665–694. https://doi.org/10.2307/3250951

Ashline, P. C., & Lai, V. S. (1995). Virtual reality: An emerging user-interface technology. *Information Systems Management*, *12*(1), 82–85. https://doi.org/10.1080/07399019508962962

Brinkman, W. P., Hoekstra, A. R. D., & van Egmond, R. (2015). The effect of 3D audio and other audio techniques on virtual reality experience. Annual Review of CyberTherapy and Telemedicine, 13, 44–48.

Bolas, M. T. (n.d.). *Human Factors in the Design of an lmmersive Display*.

Bouwels, L. (2022). *Accessible Subtitles in 6DoF 3D Virtual Reality: How Preference of Subtitle Presentation Modes relates to Presence and VR Sickness*. https://www.researchgate.net/publication/362703439

Brown, E., & Cairns, P. (n.d.). *A Grounded Investigation of Game Immersion*.
Cavalcante Raffaele, R., M Silva, F. G., & José Andrade de Carvalho, B. (2017). *Evaluation of immersive user interfaces in virtual reality first person games*. https://www.researchgate.net/publication/320387138

Cheiran, J. F. P., Bandeira, D. R., & Pimenta, M. S. (2025). Measuring the key components of the user experience in immersive virtual reality environments. *Frontiers in Virtual Reality*, *6*. https://doi.org/10.3389/frvir.2025.1585614

Chumaira, T. M., Díaz-Kommonen, L., & Bruni, L. E. (2025). Tracing affordances: mixed-methods review on techniques to study affordances in virtual reality environments. In *Cognitive Processing*. Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s10339-025-01280-x

Dillman, K. R., Mok, T. T. H., Tang, A., Oehlberg, L., & Mitchell, A. (2018). A visual interaction cue framework from video game environments for augmented reality. *Conference on Human Factors in Computing Systems - Proceedings*, *2018-April*. https://doi.org/10.1145/3173574.3173714

Cutting, J., Aung, M., Perrett, F., & Cairns, P. (2025). The development and validation of the Immersion Experience Questionnaire Short Form (IEQ-SF). *International Journal of Human Computer Studies*, *203*. https://doi.org/10.1016/j.ijhcs.2025.103580

Donald A. Norman. (n.d.). *Human action cycle*. Retrieved October 2, 2025, from https://en.wikipedia.org/wiki/Human_action_cycle

Dooley, K. (2017). Storytelling with virtual reality in 360-degrees: a new screen grammar. *Studies in Australasian Cinema*, *11*(3), 161–171. https://doi.org/10.1080/17503175.2017.1387357

Garzotto, F., Matarazzo, V., Messina, N., Gelsomini, M., & Riva, C. (2018, October 1). Improving museum accessibility through storytelling in wearable immersive virtual reality. *Proceedings of the 2018 3rd Digital Heritage International Congress, Digital Heritage 2018 - Held Jointly with the 2018 24th International Conference on Virtual Systems and Multimedia, VSMM 2018*. https://doi.org/10.1109/DigitalHeritage.2018.8810097

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). SAGE Publications.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701–721. https://doi.org/10.1037/0022-3514.79.5.701

Haggis-Burridge, M. (2020). Four categories for meaningful discussion of immersion in video games. Pure Buas. https://pure.buas.nl/ws/files/1228916/Haggis_Immersion_4CategoriesInVideoGames.pdf

Hendrix, C., & Barfield, W. (1996). The sense of presence within auditory virtual environments. Presence: Teleoperators and Virtual Environments, 5(3), 290–301. https://doi.org/10.1162/pres.1996.5.3.290

Hepperle, D., Weiß, Y., Siess, A., & Wölfel, M. (2019). 2D, 3D or speech? A case study on which user interface is preferable for what kind of object interaction in immersive virtual reality. *Computers and Graphics (Pergamon)*, *82*, 321–331. https://doi.org/10.1016/j.cag.2019.06.003

Huang, H. M. (2003). Original source for the novelty effect concept — typically cited as Huang's 2003 paper on web-based learning, though it's been extended to VR contexts. Worth verifying the exact citation through your library.

Hulusic, V., Gusia, L., Luci, N., & Smith, M. (2021). *Tangible Interfaces for VR Cultural Heritage Application-School House Virtual Museum*.

Ioanna Iacovides, Anna Cox, Richard Kennedy, Paul Cairns, and Charlene Jennett. 2015. Removing the HUD: The Impact of Non-Diegetic Game Elements and Expertise on Player Involvement. In Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15), 13–22. https://doi.org/10.1145/2793107.2793120

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

Irshad, S., & Perkis, A. (2020, October 25). Increasing User Engagement in Virtual Reality:The Role of Interactive Digital Narratives to Trigger Emotional Responses. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3419249.3421246

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (n.d.-a). *Measuring and Defining the Experience of Immersion in Games*.

John M. Flach, & John G. Holden. (n.d.). *The Reality of Experience: Gibson's Way*.

Joseph J. LaViola Jr. Ernst Kruijff Ryan P. McMahan Doug A. Bowman Ivan Poupyrev. (2017). *3D User Interfaces:Theory and Practice* (Second Edition). Pearson.

Júnior, J. R. de O., Limongi, R., Lim, W. M., Eastman, J. K., & Kumar, S. (2023). A story to sell: The influence of storytelling on consumers' purchasing behavior. In *Psychology and Marketing* (Vol. 40, Issue 2, pp. 239–261). John Wiley and Sons Inc. https://doi.org/10.1002/mar.21758

Kaur, K. (1998). *Designing Virtual Environments for Usability*.

Kemp, L. (2025, July 24). Meet VR natives, the new key demographic for virtual reality. GamesIndustry.biz. https://www.gamesindustry.biz/meet-vr-natives-the-new-key-demographic-for-virtual-reality

Kevin Brooks. (2019). There Is Nothing Virtual About Immersion - Narrative Immersion For VR and Other Interfaces. In *Motorola Labs / Human Interface Labs*.

Kuijpers, M. M., Tan, E. S., Bálint, K., & Doicaru, M. M. (n.d.). *Narrative Absorption*. https://www.ebsco.com/terms-of-use.

Kundu, R. K., Rahman, A., & Paul, S. (2021). A study on sensor system latency in vr motion sickness. *Journal of Sensor and Actuator Networks*, *10*(3). https://doi.org/10.3390/jsan10030053

Kyritsis, M., Gulliver, S. R., Morar, S., & Stevens, R. (2013). *Issues and Benefits of Using 3D Intefaces: Visual and Verbal Tasks*. www.vpython.org

Larsson, P., Väljamäe, A., Västfjäll, D., Tajadura-Jiménez, A., & Kleiner, M. (2010). Auditory-induced presence in mixed reality environments and related technology. In E. Dubois, P. Gray, & L. Nigay (Eds.), The engineering of mixed reality systems (pp. 143–163). Springer.

Lankoski, P., & Bjork, S. (2015). *Game Research Methods*. ETC Press.

Li, L., Li, M., & Pan, D. (2023). Immersive narrative design in VR game. In *Highlights in Science, Engineering and Technology AMMSAC* (Vol. 2023).

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

Lombard, M., Lee, S., Sun, W., Xu, K., & Yang, H. (2017). Presence Theory. In *The International Encyclopedia of Media Effects* (pp. 1–13). Wiley. https://doi.org/10.1002/9781118783764.wbieme0087

Longhi, R. R. (2017). Immersive narratives in web journalism. Between interfaces and virtual reality. In *Estudos em Comunicacao* (Issue 25, pp. 13–22). Universidade da Beira Interior. https://doi.org/10.20287/ec.n25.v1.a02

Markus Appel, Timo Gnambs, Tobias Richter, & Melanie Green. (n.d.). *Transportation Scale*.

Miguel-Alonso, I., Rodriguez-Garcia, B., Checa, D., & Bustillo, A. (2024). Evaluation of the novelty effect in immersive Virtual Reality learning experiences. Virtual Reality, 28, Article 27. https://doi.org/10.1007/s10055-023-00926-5

Miller, C. A., Guidry, J. P. D., Dahman, B., & Thomson, M. D. (2020). A tale of two diverse qualtrics samples: Information for online survey researchers. *Cancer Epidemiology Biomarkers and Prevention*, *29*(4), 731–735. https://doi.org/10.1158/1055-9965.EPI-19-0846

Moledina, S., Manning, C., Pobst, J., Fortier, A., Youngblood, A., Tremblay, K., & Snow, K. (n.d.). *Playful Narrative: A Toolbox for Story-Rich Mechanics The Opportunity Space: Playful Narrative 2 Existing Narrative and Mechanical Structures 2 Models for Creating Meaning 3 Divergent Meaning 3 Parallel Meaning 4 Congruent Meaning 4 Coalescent Meaning 5*.

Murray - Hamlet on the Holodeck The Future of Narrative in Cyberspace. (n.d.).

Nilsson, N. C., Nordahl, R., & Serafin, S. (2016). Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human Technology*, *12*(2), 108–134. https://doi.org/10.17011/ht/urn.201611174652

Ohn, J., And, A., Ark, M., & On -W Illiams, M. (1996). What does virtual reality NEED? : human factors issues in the design of three-dimensional computer environments. In *Int. J. Human-Computer Studies* (Vol. 44).

Pallant, J. (2020). *SPSS Survival Manual* (7th ed.). Routledge.

Pallavicini, F., Pepe, A., & Minissi, M. E. (2019). Gaming in Virtual Reality: What Changes in Terms of Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games? *Simulation and Gaming*, *50*(2), 136–159. https://doi.org/10.1177/1046878119831420

Petri, K., Feuerstein, K., Folster, S., Bariszlovich, F., & Witte, K. (2020). Effects of Age, Gender, Familiarity with the Content, and Exposure Time on Cybersickness in Immersive Head-mounted Display Based Virtual Reality. *American Journal of Biomedical Sciences*, 107–121. https://doi.org/10.5099/aj200200107

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

Pianzola, F., Riva, G., Kukkonen, K., & Mantovani, F. (2021). Presence, flow, and narrative absorption: an interdisciplinary theoretical exploration with a new spatiotemporal integrated model based on predictive processing. *Open Research Europe*, *1*, 28. https://doi.org/10.12688/openreseurope.13193.1

Procci, K., & Bowers, C. (2011). An examination of flow and immersion in games. *Proceedings of the Human Factors and Ergonomics Society*, 2183–2187. https://doi.org/10.1177/1071181311551455

P. Salomoni, C. Prandi, M. Roccetti, L. Casanova, and L. Marchetti. 2016. Assessing the efficacy of a diegetic game interface with Oculus Rift. In 2016 13th IEEE Annual Consumer Communications Conference (CCNC), https://doi.org/10.1109/CCNC.2016.7444811

Qin, H., Patrick Rau, P. L., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human-Computer Interaction*, *25*(2), 107–133. https://doi.org/10.1080/10447310802546732

Raffaele, R. C., & Fernandes, F. (2016). *Virtual Reality: an analysis of Immersive UI for first person games*. https://www.researchgate.net/publication/314183439

Raffaele, R., Carvalho, B., Lins, A., Marques, L., & Soares, M. M. (2016). Digital game for teaching and learning: An analysis of usability and experience of educational games. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9747*, 303–310. https://doi.org/10.1007/978-3-319-40355-7_29

Rebecca Noll. (2021). *Can't touch this? Exploring The Concept Of Affordance In VR*.

Riecke, B. E., LaViola, J. J., & Kruijff, E. (2018). *3D user interfaces for virtual reality and games*. 1–94. https://doi.org/10.1145/3214834.3214869

Roth, C., & Koenitz, H. (2016). Evaluating the user experience of interactive digital narrative. *AltMM 2016 - Proceedings of the 1st International Workshop on Multimedia Alternate Realities, Co-Located with ACM Multimedia 2016*, 31–36. https://doi.org/10.1145/2983298.2983302

Sala, N. (2020). *Virtual Reality, Augmented Reality, and Mixed Reality in Education* (pp. 48–73). https://doi.org/10.4018/978-1-7998-4960-5.ch003

Schleußinger, M. (2021). Information retrieval interfaces in virtual reality - A scoping review focused on current generation technology. *PLoS ONE*, *16*(2 February). https://doi.org/10.1371/journal.pone.0246398

Seah, M., & Cairns, P. (2008). From immersion to addiction in videogames. *Electronic Workshops in Computing*. https://doi.org/10.14236/ewic/hci2008.6

Influence of User Interface markers design on player's immersion into a Virtual Reality game narrative. // Marking tangable user interface in VR games to support narrative immersion.

Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L., & Wolman, I. (2019). Framework of Affordances for Virtual Reality and Augmented Reality. *Journal of Management Information Systems*, *36*(3), 683–729. https://doi.org/10.1080/07421222.2019.1628877

Sutcliffe, A. G., & Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. *Behaviour and Information Technology*, *19*(6), 415–426. https://doi.org/10.1080/014492900750052679

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.

Toyama, T., Sonntag, D., Orlosky, J., & Kiyokawa, K. (2015). Attention engagement and cognitive state analysis for augmented reality text display functions. *International Conference on Intelligent User Interfaces, Proceedings IUI*, *2015-January*, 322–332. https://doi.org/10.1145/2678025.2701384

Triberti, S., Sapone, C., & Riva, G. (2025). Being there but where? Sense of presence theory for virtual reality applications. In *Humanities and Social Sciences Communications* (Vol. 12, Issue 1). Springer Nature. https://doi.org/10.1057/s41599-025-04380-3

Types of UI in Gaming: Diegetic, Non-Diegetic, Spatial and Meta. (n.d.). Retrieved October 2, 2025, from https://medium.com/@lorenzoardeni/types-of-ui-in-gaming-diegetic-non-diegetic-spatial-and-meta-5024ce6362d0

U. A. Chattha, U. I. Janjua, F. Anwar, T. M. Madni, M. F. Cheema and S. I. Janjua, "Motion Sickness in Virtual Reality: An Empirical Evaluation," in IEEE Access, vol. 8, pp. 130486-130499, 2020, doi: 10.1109/ACCESS.2020.3007076.

Weger, U. W., Loughnan, S., Sharma, D., & Gonidis, L. (2015). Virtually compliant: Immersive video gaming increases conformity to false computer judgments. *Psychonomic Bulletin and Review*, *22*(4), 1111–1116. https://doi.org/10.3758/s13423-014-0778-z

Wolfe, A., Louchart, S., & Loranger, B. (2022). The Impacts of Design Elements in Interactive Storytelling in VR on Emotion, Mood, and Self-reflection. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13762 LNCS*, 616–633. https://doi.org/10.1007/978-3-031-22298-6_40

Yang, Q., Hua, S., & Dai, G. (2024). Design of Visual Communication Experience Based on Virtual Reality Technology. In *J. Electrical Systems* (Vol. 20, Issue 3).